ABSTRACT

The use of Stereotactic Radiation Therapy (SRT) employing one large fraction of radiation, as in stereotactic radiosurgery (SRS), or few fractions of high doses, has continuously increased due to the technical development and the progress in dose delivery complemented by the positive clinical experience. The success of stereotactic radiation therapy depends on many clinical, dosimetric and radiobiological factors. For SRS in particular, the delivery of a highly conformal dose distribution to the target in one fraction allowing at the same time the sparing of the normal tissue and the critical structures is part of the basic concept of the technique. Provided that the highly accurate radiosurgical equipment available today is used, planning and delivering the prescribed dose distribution is an achievable goal, and therefore the main issue to be solved is the definition of the target. As the target volume in radiosurgery is usually defined without margins, the success of the stereotactic approach critically depends on the accurate delineation of the target which could be identified as a factor of key importance. In addition, the delineation of the Organs At Risk (OAR) is also critical.

The purpose of this work was to evaluate the current degree of variability for target and OAR contouring and to establish methods for analysing multi-observer data regarding structure delineation variability.

A multi-center target and OAR delineation study was initiated. Two complex and six common cases to be treated with SRS were selected and subsequently distributed to centers around the world performing Gamma Knife® radiosurgery for delineation and treatment planning. The resulting treatment plans and the corresponding delineated structures were collected and analysed.

Results showed a very high variability in contouring for four complex radiosurgery targets. Similar results indicating high variability in delineating the OAR and reporting the doses delivered to them were also reported. For the common radiosurgery targets however, a higher agreement in the delineation was observed, although lower than expected.

The assessment of the quality of treatment planning for radiosurgery is usually performed with respect to the coverage of the target, the planning specificity, and dose to the sensitive structures and organs close to the target. However, physical dose conformity to the target does not guarantee the success of the treatment. The
assessment of the plan quality should also be performed with respect to the clinical outcome expressed as probability of controlling the target that should be irradiated. In this respect, this study also aimed to create the framework for assessing the impact of the inaccuracy in delineating the target on the predicted treatment outcome for radiosurgery targets known for their high potential to invade the neighbouring normal tissue, using radiobiological models. In addition, radiobiological models have also been used to determine the tumour control probability accounting for the oxygenation for stereotactic radiation therapy targets.

The results suggest that radiobiological modelling has the potential to add to the current knowledge in SRS by theoretically assessing the key factors that might influence the treatment outcome.
LIST OF PAPERS

The following papers are included in the thesis. Reprints were made with permission from the publishers.

PAPER I: Variability in target delineation for cavernous sinus meningioma and anaplastic astrocytoma in stereotactic radiosurgery with Leksell Gamma Knife Perfexion

DOI: 10.1007/s00701-014-2235-1

PAPER II: A multi-institutional study of the variability in target delineation for six targets commonly treated with radiosurgery

H. Sandström, H. Jokura, I. Toma-Dasu, Manuscript

PAPER III: Preliminary assessment of Organs-at-Risk contouring practices in radiosurgery institutions around the world


PAPER IV: Radiobiological framework for the evaluation of stereotactic radiosurgery plans for invasive brain tumors

DOI: 10.1155/2013/527251

PAPER V: To fractionate or not to fractionate? That is the question for the radiosurgery of hypoxic tumors

DOI: 10.3171/2014.8.GKS141461
## CONTENTS

ABSTRACT ........................................................................................................................ 3

LIST OF PAPERS ............................................................................................................. 5

ABBREVIATIONS ............................................................................................................ 9

1. INTRODUCTION ...................................................................................................... 11

2. GAMMA KNIFE RADIOSURGERY ...................................................................... 13

   2.1 Target and organs at risk contouring ................................................................. 14

   2.2 Treatment plans evaluation parameters .............................................................. 16

   2.3 Single fraction or fractionated stereotactic radiosurgery .................................... 18

3. VARIABILITY IN CONTOURING AND THE SEARCH OF THE TRUE VOLUME .......................................................................................................................... 20

   3.1 DICOM files in the Leksell GammaPlan ............................................................... 20

   3.2 Analysis methodology of multicenter data ......................................................... 23

   3.5 Pitfalls and solutions for determining the true volume ...................................... 29

4. SUMMARY OF PAPERS ......................................................................................... 32

5. OUTLOOK .................................................................................................................. 34

REFERENCES ................................................................................................................ 35
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AI</td>
<td>Angiography Imaging</td>
</tr>
<tr>
<td>AV&lt;sub&gt;x&lt;/sub&gt;</td>
<td>Agreement Volume of X % concordance</td>
</tr>
<tr>
<td>AVI</td>
<td>Agreement Volume Index</td>
</tr>
<tr>
<td>AVM</td>
<td>Arteriovenous Malformation</td>
</tr>
<tr>
<td>CBCT</td>
<td>Cone-Beam Computed Tomography</td>
</tr>
<tr>
<td>CNS</td>
<td>Central Nervous System</td>
</tr>
<tr>
<td>COM</td>
<td>Center of Mass</td>
</tr>
<tr>
<td>CT</td>
<td>Computed Tomography</td>
</tr>
<tr>
<td>DVH</td>
<td>Dose Volume Histogram</td>
</tr>
<tr>
<td>LGP</td>
<td>Leksell Gammaplan</td>
</tr>
<tr>
<td>LQ</td>
<td>Linear Quadratic</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnetic Resonance Imaging</td>
</tr>
<tr>
<td>OAR</td>
<td>Organs at Risk</td>
</tr>
<tr>
<td>PET</td>
<td>Positron Emission Tomography</td>
</tr>
<tr>
<td>ROI</td>
<td>Region of Interest</td>
</tr>
<tr>
<td>RS</td>
<td>Radiosurgery</td>
</tr>
<tr>
<td>RT</td>
<td>Radiotherapy</td>
</tr>
<tr>
<td>SBRT</td>
<td>Stereotactic Body Radiotherapy</td>
</tr>
<tr>
<td>SRS</td>
<td>Stereotactic Radiosurgery</td>
</tr>
<tr>
<td>TCP</td>
<td>Tumour Control Probability</td>
</tr>
<tr>
<td>TPS</td>
<td>Treatment Planning System</td>
</tr>
<tr>
<td>STAPLE</td>
<td>Simultaneous Truth and Performance Level Estimation</td>
</tr>
</tbody>
</table>
1. Introduction

For more than 40 years, intracranial stereotactic radiosurgery (SRS) has been used for the management of malignant and benign lesions as well as for functional disorders in the brain. The first commercial Gamma Knife, Model U, was installed at University of Pittsburgh in 1987 and by the year of 2011, nearly 676,000 patients have been treated worldwide with the Gamma Knife (Wu et al. 1990, Leksell Gamma Knife Society 2011). The advantage of SRS in comparison to external beam linac-based radiotherapy (RT) is the application of multiple focused radiation beams achieving high normal tissue sparing at the same time with delivering a high dose to the target in one fraction delivered with submillimeter precision (Metha et al. 2005). Therefore the target definition and subsequently the corresponding treatment planning are very important in SRS since despite the high conformity and normal tissue sparing in SRS, an inaccuracy in target definition could negatively impact the tumour control.

SRS dose distributions conform to the contoured target with high dose gradients outside the target, allowing thus the possibility of delivering a high dose to the target and restricting the doses to Organs at Risk (OAR) in the proximity of target (Leksell 1951, Mack et al. 2002). The roles of the experienced neurosurgeon, physicist and radiologist are central in the process of achieving the overall aim of the treatment with respect to the normal tissue dose restrictions, conformity parameters for the plan and patient comfort. A review of normal tissue dose restrictions has been provided in the Quantitative Analysis of Normal Tissue Effects in the Clinic (QUANTEC) studies (Lawrence et al. 2010, Mayo et al. 2010) highlighting the lack of data and the variability in reporting doses to OAR in SRS planning. This emphasizes the importance of Central Nervous System (CNS) and brain OAR contouring guidelines for SRS treatments.

The high accuracy of the SRS techniques available today, especially Gamma Knife® radiosurgery (GKRS), is not used at its full potential because of the limited accuracy in contouring the targets and the healthy structures. There is a large number of studies in the literature presenting contouring variability for targets and OAR outside the brain (Berthelet et al. 2002, Giraud et al. 2002, Weiss et al. 2003, Breen et al. 2007, Tyng et al. 2009, White et al. 2009, van Mourik et al. 2010, Genovesi et al. 2011, Feng et al. 2012, Cui et al. 2015). However, the literature on target contouring variability specifically for the brain region and SRS, is limited to few studies which all present variability results much higher than the accuracy of the technique itself with respect to the delivery of the dose with high conformity (Buis et al. 2005, Yamazaki et al. 2011,
Stanley et al. 2013, Sandström et al. 2014 – **Paper I**, Sandström et al. 2015a – **Paper III**, Sandström et al. 2015b – **Paper II). Again, this emphasizes the importance of the contouring process in the sense that the variability could shadow the accuracy of the SRS technique. It has been discussed that target and OAR definition and contouring is a source of uncertainty in RT in general, additional to target motion and patient setup (Weiss and Hess 2003, Rasch et al. 2005, Steenbakkers et al. 2006, Njeh 2008). Guidelines for OAR contouring and dose constraints have been reported for brain tumours planned for fractionated RT but no such recommendation exists for SRS targets (Scoccianti et al. 2015). To be able to confirm the need for contouring guidelines, studies on contouring variability are necessary as shown in the work by Sandström et al. (2015a) – **Paper III** in which it was reported that the variability in OAR contouring is high when no guidelines are provided to participants. The need for guidelines was also shown in other studies for other tumour sites (Riegel et al. 2006, Castro Pena et al. 2009, Genovesi et al. 2011). It is therefore important to be able to identify the factors contributing to the variability in target and OAR contouring to be able to minimise their influence. The choice of imaging methods and the experience and training of the practitioner performing the delineation, are some of the possible factors influencing the contouring methodology which in turn might impact the resulting contours.

Assessing the variability in the delineation of the target should also be complemented by attempting to determine the most probable target, the *true target* volume, for SRS.

In the era of advanced development of methods for single fraction and hypofractionated stereotactic radiosurgery, the quest for the *true target* volume setting the foundation for contouring practice recommendations is therefore of key importance.
2. Gamma Knife Radiosurgery

Gamma Knife® Perfexion™ using 192 Cobalt-60 sources is commercially available since 2006 (Lindquist and Paddick 2007, Novotny et al. 2008). Cobalt-60 undergoes beta decay with a half-life of 5.27 years and an effective average energy of 1.25 MeV.

The sources, with an initial activity of approximately 1.11 TBq, are arranged in a semi-spherical pattern, alongside eight moveable sectors, with beams intersecting at the isocenter. The movement of the sources allows three possible collimator positions and one blocked position for each of the sources (Lindquist and Paddick 2007).

Images from various imaging techniques such as, magnetic resonance imaging (MRI), computed tomography (CT) and positron emission tomography (PET) as well as planar images from angiography (AI) are supported by the treatment planning system (TPS) Leksell GammaPlan® (Elekta Instrument AB, Stockholm, Sweden) and used in the treatment planning process.

Accurate positioning is ensured by the use of the Leksell® Coordinate Frame. The Leksell® Coordinate System is correlated to the Leksell® Coordinate Frame and it provides the Cartesian X, Y and Z coordinates in Leksell® space. The stereotactic frame is mounted on the head of the patient prior to treatment and the patient is imaged with the frame attached which constitutes the reference image and coordinate system for treatment planning. This image is co-registered to other tomographic or planar morphological or functional images required for treatment planning. Furthermore, since the frame is rigid and attached with screws to the patients head, it prevents the movement of the patient relative to the coordinate system.

The dose distribution resulting from one beam configuration in the Gamma Knife® Perfexion™ system is generally spherical in shape. Small targets can be treated with a single so called "shot", corresponding to one beam configuration while larger or more complicated lesions are treated with multiple shots. The overlap of shots results in a non-spherical dose distribution which can be conformably adjusted to the target. Individual shots can furthermore be angularly weighted to create non-spherical dose distributions, for each shot, which adds to the possibility of creating a complex treatment plan. An individual shot blocked in one or several directions, a hybrid shot, facilitates dose sparing of adjacent normal tissue which is beneficial for cases with OAR in the proximity of target (Lindquist and Paddick 2007, Petti et al. 2008). The use of multiple shots, and shots of smaller collimator sizes, leads however to an increase in treatment time. Users need consequently to keep this factor in mind and create the
optimal treatment plan with respect to both target obliteration and normal tissue sparing as well as reduced treatment time.

2.1 Target and organs at risk contouring

The target volume, the volume that will receive the prescribed dose, could be manually contoured as well as contoured with the assistance of automatic segmentation. The target could be a tumour, thus a malignant pathology, a benign lesion or a functional target. As examples of radiosurgery targets one could mention metastases, cavernous sinus meningiomas, vestibular schwannomas, trigeminal neuralgia or arteriovenous malformations (AVM) respectively. Radiosurgery, as quoted from neurosurgeons, implies that no margin is applied to the contoured target.

In the treatment planning for SRS targets, implying a single- or hypo-fractionated regime, the planner does always optimize the treatment based on the assumption that all cells that should be eradicated are included in the target of interest with maximal sparing of surrounding normal tissue. The treatment planning is frequently dependent on the choice of images and MRI images are regularly used. Without knowledge of target pathology, the target contouring depends mainly on image contrast, user knowledge and experience. Thus, for particular types of radiosurgery targets, namely malignant lesions with high potential for invading the surrounding normal tissue like astrocytomas, those that are not treated with linac-based radiotherapy, the delineation of the target should account for their invasive character.

Sandström et al. (2013) – Paper IV, reported in a modelling paper the rapid decrease in the probability of controlling the tumour when target cells are exponentially distributed within the outside rim of contoured target, decreasing with the distance to target boundary. The study has therefore highlighted the importance of accurate contouring depending on the pathology of the target.

Contouring of OAR is obviously necessary for the evaluation of the doses delivered to the OAR in relation to the tolerance doses for the same organs. In SRS of the brain, there is no consensus regarding the tolerance doses for relevant OAR. Therefore, an OAR Standardization Working Group supported by the Leksell Gamma Knife Society was established (Chung et al. 2015). According to the OAR Standardization Working Group the reported range of accepted tolerance doses for OAR in the brain, is large. Another issue brought up by the OAR Standardization Working Group is the lack of consensus on the use of particular images, produced by specific imaging modalities, for OAR contouring. As an illustration of the inconsistency in the use of a particular
image type for delineating an OAR, Figure 1 shows the cochlea contoured by six Gamma Knife users in a vestibular schwannoma case. All panels in Figure 1 correspond to the same slice (Sandström et al. 2015a - Paper III). The obvious differences in the size and shape of cochlea in this slide raise the concern regarding the need for recommendations of particular images for specific structures.

Figure 1. Six images of the same slice of cochlea and the corresponding contour delineated on the CT image (A and B), fused image of CT and T2-weighted MRI (C) and T2-weighted MRI (D-F). Bottom figure shows the overlapping contours.
2.2 Treatment plans evaluation parameters

Several parameters used for the evaluation of the plans are considered at the treatment planning stage. The coverage index (value between 0-100%) is the ratio of the target volume receiving the prescribed dose to the whole volume of the target. A generally accepted value of the coverage index is ≥95% (Lomax and Scheib 2003) which accounts for the possibility to lower the target dose in the direction of an OAR. The selectivity index, with an identical range as the coverage index, provides a measure of prescribed dose “spilling over” to the normal tissue surrounding the contoured target. The accepted value for the selectivity index varies and it is reported to be dependent on size and shape of the target; larger spherical targets should be planned with a selectivity equal or larger than 85% while plans for small or irregular shaped targets could be accepted with a selectivity index equal or larger than 50% or 70%, respectively. Conformity index (CI) and Paddick conformity index (PCI) are two other indexes describing the conformity of a treatment plan relative to the contoured target volume (Shaw et al. 1993, Paddick 2000). The conformity index, equation 2.1, is the ratio of the prescription isodose volume, \( V_p \), and the target volume, \( V_T \), respectively and is greater or equal to one. Paddick conformity index is the product of the coverage and selectivity, equation 2.2, where \( V_{TP} \) is the target volume covered by the prescribed isodose. The PCI takes into account both over- and under-treatment with respect to the delineated target volume.

One important aspect in the evaluation and acceptance of a treatment plan is the voxel sizes of the volumes used for planning. A small target would benefit from smaller sizes of the voxels for a clear representation of the indices due to the voxel coverage close to structure boundaries.

\[
CI = \frac{V_p}{V_T} \tag{2.1}
\]

\[
PCI = \frac{V_{TP}}{V_T} \times \frac{V_{TP}}{V_p} \tag{2.2}
\]

The dose fall-off outside the treatment volume is described using the gradient index (GI) (Paddick and Lippitz 2006) which is defined as the ratio of volume of half the prescribed isodose to the prescribed isodose volume, see equation 2.3.
\[ GI = \frac{V_{P/2}}{V_P} \]

Dose Volume Histograms (DVHs) are also used to evaluate the generated plans in terms of OAR doses and target coverage. Examples are shown in Figure 2 which displays 12 DVHs corresponding to plans made by 12 Gamma Knife centres from around the world for a cavernous sinus meningioma together with OAR DVHs for the optic chiasm in the same plans.

![Figure 2. Dose Volume Histograms for cavernous sinus meningioma including all contoured structures for target as well as the chiasm as one of the organs at risk of interest. The relative volume in percent is plotted as a function of dose (Gy). The plans were made by 12 Gamma Knife centers from around the world. All plans were made for the 12-16 Gy dose to the 50% isodose.](image)

As previously mentioned, another parameter of concern in the evaluation of SRS plans is treatment time. As the source decays the treatment delivery could be prolonged and a complicated treatment plan with many shots will result in an extended treatment time. The treatment time is additionally dependent on the equipment. However, Regis et al. (2009) reported a reduced treatment time, from 65 min to less than 45 min, comparing Gamma Knife® model 4C to Perfexion™.
2.3 Single fraction or fractionated stereotactic radiosurgery

Radiosurgery of targets in the brain has for decades been defined as one fractional treatment, analogous to surgery and hence the way the concept of radiosurgery was introduced. Alternative technologies have recently emerged which allow for fractionated treatments of lesions in the brain with equipment initially developed for single fraction treatments. In a patient perspective the benefit is high since the technology reduces the invasive component of frame-placement. Stereotactic techniques have as well been developed for extracranial targets, such as stereotactic body radiotherapy (SBRT), where a conformal dose is delivered in one or few fractions. The main difference and therefore limitations of one over the other, comparing radiosurgery of the brain to SBRT, is organ movement which is solved by the application of margins in SBRT (Kavanagh et al. 2006, Potters et al. 2010). The main reason for a fractionated regime is to take advantage of the difference in radiosensitivity of pathologic tissue compared to normal tissues as well as allowing for reoxygenation of target tissue between fractions. A fractionated SRS regime has been shown beneficial for patients with lesions close to the optic pathways (Kim et al. 1999, Kim et al 2008, Jee et al. 2014, Devriendt et al. 2015). The inverse correlation between the volume of a target and the maximum dose that can be prescribed, with respect to normal tissue toxicity, can be overcome by a hypo-fractionated treatment regime.

A number of studies have shown that adequate single fraction SRS treatments can be performed with doses between 15-24 Gy for metastases which are malignant lesions, and 12-13 Gy and 14-16 Gy for benign lesions such as vestibular schwannomas and meningiomas respectively (Stafford et al. 2001, Sneed et al. 2002, Hasegawa et al. 2005, Mehta et al. 2005, Chopra et al. 2007, Kondziolka et al. 2008, Koyfman et al. 2010, Suh JH 2010). Control rates for common benign intra-cranial tumours treated with SRS have proven to be quite high (Pollock 2009). However, studies on fractionated regimes for radiosurgery are limited. Toma-Dasu et al. (2014) – Paper V, showed in a modelling study that large metastases with a known relatively high hypoxic fraction of cells might benefit from a fractionated regime while smaller moderately hypoxic metastatic tumours do not benefit to the same extent.

Several Gamma Knife® techniques are available for fractionated high-dose radiosurgical treatments. The purpose is either the fractionation of a therapeutic dose for one target or the treatment of numerous targets, such as multiple metastases. The eXtend™ System is a fully integrated addition to the Leksell Gamma Knife® perfexion™ and allows fixation through a vacuum mouthpiece, therefore enabling the non-invasive fractionation of treatments (Sayer et al. 2011, Schlesinger et al. 2012).
The treatment safety measures are ensured by the rapid shut down and retracement of sources in case of vacuum discharge. The newly developed version of the Gamma Knife®, the Icon™, has moved cranial radiosurgery from a frame-based to a frameless approach. This model incorporates a cone-beam CT (CBCT) for the definition of stereotactic coordinates together with motion detection by optical tracking of fiducial markers on patient's nose. Hypo-fractionated treatments of larger lesions or lesions extending through radiosensitive tissues are now available with higher patient comfort. This will further merge the radiosurgery community with conventional radiotherapy.
3. Variability in contouring and the search of the true volume

With the high-dose conformity techniques available today, accurate contouring of the structures of interest is especially important. Several studies have shown variation in the contouring for targets outside the brain (Villeirs et al. 2005, Jeanneret-Sozzi et al. 2006, Yamazaki et al. 2007) as well as for cranial targets (Weltens et al. 2001). Anatomical and functional images are only representations of the normal tissue and pathological changes and the true size, shape and location of the target lesion may not be known by the person contouring it. Therefore, any clinically contoured volume could be regarded as the best estimate of a true volume with respect to tumour control and normal tissue complications, based on the assumption that all contours are drawn by specialists within the field. The choice of images used for contouring of both target and OAR should moreover be made according to the optimal visualization of the structures. The question then is: can several volumes of the same structure contoured by different practitioners be used to estimate the true target?

Several methods trying to quantify the inter-observer contouring inconsistency have been described in the literature for several targets and the majority deal with overlapping volumes and indices derived from them (Logue et al. 1998, Struikmans et al. 2005, Song et al. 2006, Landis et al. 2007, Petersen et al. 2007, Li et al. 2009, Vorwerk et al. 2009, Rasch et al. 2010, Lütgendorf-Caucig et al. 2011, Altorjai et al. 2012, Sandström et al. 2014 – Paper I, Sandström et al. 2015a – Paper III, Sandström et al. 2015b – Paper II). Sandström et al. (2014) - Paper I proposed a method that derives the average target based on an agreement matrix where each voxel has a value between 0-N where N is the total number of segmentations (available contours) for that particular target. This method is not limited to the number of structures analysed and furthermore, does not limit the user to calculating only the average target as the agreement matrix can be segmented based on the level of agreement of choice.

3.1 DICOM files in the Leksell GammaPlan

The Leksell GammaPlan® (LGP) treatment planning system supports images in the Digital Imaging and Communications in Medicine (DICOM) format. DICOM is a valuable tool that gives the user the possibility of exporting the data
regarding the images and contouring from the TPS to other software platforms providing the option of performing additional analysis of the data outside the LGP.

The DICOM export from LGP generates a set of DICOM files: structure set files (RTSS), images (RTImage), treatment plan files (RTPlan) and dose matrices (RTDose). These can be imported in the MATLAB software (MathWorks, Inc., Natick, Massachusetts, USA).

The RTSS-file provides spatial information, i.e. the coordinates for the contoured regions of interest (ROIs), for each of the structures contoured in the plan where the resolution in X and Y is determined in the export process and the Z resolution is determined by the image sequence/modality. The RTImage-file is the set of images, in the TPS, which the user chooses to export. The RTPlan-file is generated at each export and gives the treatment planning parameters such as user name, patient name, type of radiation, patient setup, number of fractions and gantry angle. The RTDose-files give the dose matrix for the entire skull, the target volume and for the structure/structures exported.

An illustration of the DICOM information provided and used for volumetric comparison of the targets and the corresponding dose distributions, is shown in Figure 3.

Figure 3. Illustration of the structure of DICOM RTSS files with sub-levels for volumetric analysis of contours.
The RTSS-file provides information regarding the images used for contouring, name and number of each specific structure, if an automatic tool or manual contouring has been used for contouring, the contour data grouped by slice and the actual coordinates of the contours. At this level the individual volumes can be visualised as 3D volumes as illustrated in Figure 4 for a case of cavernous sinus meningioma and pituitary adenoma.

![Figure 4. Example illustrations of the 3D ROIs for cavernous sinus meningioma (left) and pituitary adenoma, (right) retrieved from the contour data.](image)

To be able to handle and compare regions of interest, a built-in function in MATLAB can be used which creates a binary matrix as illustrated in Figure 5. Both the dose matrices and structure matrices are based on the same coordinate system, the Leksell space, and can therefore be analysed within the identical frame of reference.

![Figure 5. Illustration of five slices of the binary matrix of a cavernous sinus meningioma where voxels with the value of 0 (white) and 1 (black) correspond to target and normal tissue respectively.](image)
Figure 6 illustrates the prescribed dose distribution compared to the delineated structure, as isosurfaces, for one case of cavernous sinus meningioma. In the header of the RTDose file, a 3-dimensional starting point can be extracted which provides the basis of both target and dose matrix placement. From this, the coordinate system, identical in X, Y and Z, for the target and dose matrix can be generated. This facilitates the possibility of comparing and/or overlaying structures with dose distributions and calculating the conformity indices.

![Figure 6. 3D surface plots of the structure matrix (blue) compared to the prescribed dose (green) for a case of cavernous sinus meningioma.](image)

3.2 Analysis methodology of multicenter data

Several methods can be used to visualize and quantify the variability in contouring of targets or OAR (Sandström et al. 2014 – **Paper I**). Determining the agreement volumes has both a qualitative and a quantitative side and it takes into account the variability in size, shape and position of the target or OAR on voxel-level. The method proposed by Sandström et al. (2014) – **Paper I** generates a 3D agreement map of overlapping structures where the voxels have a value between 0-N, N being the number of overlapping structures. The values between zero and N are a measure of diverse levels of concordance (Sandström et al. 2014 – **Paper I**). From these maps, volumes of interest and their agreement level can be determined and compared, visually and quantitatively. An example of how the agreement levels can be visualized
is displayed in Figure 7, where isolines show different levels of agreement (values 1-11 in the example provided). A region of higher value indicates a higher agreement between the delineated structures.

![Figure 7. Example of agreement levels for cavernous sinus meningioma (top) and pituitary adenoma (bottom). The colour bar on the right gives the value of agreement. Blue corresponds to the highest agreement and outer levels of green correspond to higher disagreement. The outermost region corresponds to the union of all contoured structures while the innermost line or region is the union of contours.](image)

Another way of displaying the agreement volumes is by plotting them as slices, as in Figure 8. In this figure, the agreement volumes for a case of cavernous sinus meningioma and pituitary adenoma are illustrated by surfaces corresponding to different agreement levels. The colour bar on the right indicates the level, from 0 (complete agreement) to N (number of overlapping contours up to complete agreement).
agreement). The area with the value of 1 or N is the surface of the encompassing volume (union volume) or common volume (intersect volume), respectively.

Figure 8. Illustration of agreement volumes for a case of cavernous sinus meningioma (top) and pituitary adenoma (bottom) for six slides of the agreement volume. Darker shades correspond to lower agreement while both black and white correspond to complete agreement.

These measures have been used to quantify the variability in structure contouring and can be used to compare two or more contoured structures (Rasch et al. 1999, Weltens et al. 2001). Although widely used in the literature, the downside of these measures is the dependence on the number of participants in the study and hence the number of analysed structures. The common volume can remain unchanged but by no means become larger with additional structures added. By adding a volume to the existing
analysis the resulting common volume will remain the same, if the structure completely encompasses the previous common volume, or lower the common volume in the case of disagreement with the previous analysis. The main limitation of this method is therefore the restriction of inter-target comparison.

Sandström et al. (2014) – Paper I presented another method for determining and displaying the contouring variability, illustrated in Figure 9 as spherical isosurfaces (the common volume, average volume and encompassing volume) compared to the actual volumes. These figures show the spherical representation of a specific volume of any shape, in this example a cavernous sinus meningioma in three and two dimensions.

An index to quantify the contouring variability by comparing targets to each other, is the Agreement Volume Index (AVI) which has been described in the literature under different nomenclature but with the same method for calculating it (Rasch et al. 1997, Yamamoto et al. 1999, Fox et al. 2005, Struikmans et al. 2005, Voroney et al. 2006,
AVI is the ratio of common to encompassing volume and has an ideal value of 1. Similar to the common and encompassing volume, this index is dependent on the number of participants in a contouring study.

With the assumption that a best estimate, the true volume, of a specific structure is known, each additional delineated volume could be analysed relative to the actual one in terms of accuracy and precision. The accuracy is the level of correspondence to the true volume (0-1) (sensitivity) and the precision is the level of defining only the true volume (0-1) (specificity). STAPLE (Simultaneous Truth and Performance Level Estimation) is an iterative algorithm which takes the analysis further and generates a probabilistic most likely segmentation of a volume based on numerous segmentations. Moreover, it calculates the sensitivity and specificity of each segmentation input (Warfield et al. 2004) and uses these values in an iterative approach; iteration n uses the sensitivity(n-1) and specificity(n-1). The iterative approach reaches a level of stability, reportedly before the 20th iteration, at which the most probable segmentation and the specificity and sensitivity for each segmentation input can be extracted. In Figure 10, panel A, the true volume determined using the STAPLE method is compared to the 50% agreement volume determined applying the method described in Sandström et al (2014) - Paper I, illustrated as 3D surfaces. Panel B illustrates the overlapping contours of the 50% agreement volume and the true volume. The regions where the two volumes do not coincide are highlighted in red.

Another way of analysing multicenter contouring data is the assessment of the position of the center of mass (COM) which is the weighted point in space corresponding to the center of mass of that volume. The purpose of incorporating this measure in the analysis of multi-observer data is to add a complementing criterion to the analysis of target delineation variability by investigating how the coordinates of COM of contoured volumes differ for different volumes. Examples of the position of the COM together with the common and encompassing volume are illustrated in Figure 11 for a cavernous sinus meningioma (A) and pituitary adenoma (B).
Figure 10. (A) Comparison of the 50% agreement volume (AV50, blue) and the STAPLE generated true volume (green). (B) Overlapping contours for the 50% agreement volume and true volume where red areas correspond to non-agreement between contours.

Figure 11. Illustration of the center of mass (red points) together with the common volume (blue) and encompassing volume (green) for (A) cavernous sinus meningioma and (B) pituitary adenoma.
3.5 Pitfalls and solutions for determining the true volume

The accuracy in determining the true volume of a given structure relies on the expertise of the observers performing the segmentations. The experience and knowledge of the individual participants in a segmentation study are not usually evaluated and therefore it is important to minimise the impact of these factors, or conclude that they are not relevant.

Figure 12 shows the sensitivity (accuracy), generated with the STAPLE algorithm as function of experience in years for nine individual participants in the segmentations study for a cavernous sinus meningioma target. With respect to the true volume, the sensitivity varies between 0.58 and 0.90. The interpretation of this is that observers contour this target with accuracy in covering the true volume between 58% and 90%. Based on these results, the influence of experience and training should be further investigated. The STAPLE algorithm is furthermore sensitive with respect to the excess volume outside the analysed structures, i.e. the volume representing the normal tissue.

Figure 12. Segmentation sensitivity as a function of experience (years) for a case of cavernous sinus meningioma generated with the STAPLE algorithm.

Figure 13 shows the sensitivity and specificity as a function of excess volume for a metastasis case and clearly indicates how the specificity is improved as the normal tissue volume increases. This is the result of the fact that the specificity is a relative number with respect to the amount of normal tissue spared when each segmentation input is compared to the true volume and the recommended calculation should be performed with a matrix conformal to the contoured structures.
In the construction of a study on contouring variability, many parameters and influencing factors have to be taken into account. The initial instructions to participants have to be minimally influencing the contouring work and simultaneously enough elaborated to eliminate the external influencing factors such as data interpolation, proper imaging used and suitable targets.

Sandström et al. (2015a) - Paper III reported results on a study for OAR where the majority of planning parameters were non-fixed. The results showing the OAR delineation variability was highly influenced by the planning methodology of the participants although reflecting a clinical reality. In another study by Sandström et al. (2015b) - Paper II, where targets regarded as common for SRS were chosen for delineation, the variability was, as expected, lower. In that study based on a set of common targets all influencing factors, except experience and imaging choice, were eliminated.

Contouring results for the OAR showed a high variability in the volumes contoured, the extent of volumes contoured, terminology and imaging modalities chosen for delineation (Sandström et al. 2015a - Paper III). Therefore, it could be suggested that a study on delineation of the normal organs and tissues needs to include specific and detailed instructions for the participants since the most contributing factor to the OAR contouring variability was in the view on which part of a structure to contour. The
OAR Standardization Working Group aims of establishing contouring recommendations for radiosurgical targets. This concerns the accepted tolerance dose, imaging modalities, terminology and OAR structures of relevance for specific targets. With these recommendations at hand, a robust study of the OAR contouring can be initiated.
4. Summary of papers

Target and OAR contouring variability has been established in this thesis to be one of the key issues in SRS. A variety of issues have been explored in this project, from quantifying target and OAR variability in delineation to the radiobiological framework of analysing the consequences of this variability. Reducing delineation variability can be done by establishing contouring guidelines for both targets and OAR based on segmentation atlases for OAR and recommendations for targets in terms of the use of imaging modalities, prescription doses, use of margins depending on the pathology, etc.

Paper I

The aim of the study was to quantify the variability in target delineation for two complex SRS targets: one cavernous sinus meningioma imaged with MRI and CT and one anaplastic astrocytoma imaged with MRI and PET. Additionally, the study aimed to investigate the dosimetric implications of variability in target delineation with respect to the plan conformity. Twenty centres chosen for their experience with Leksell Gamma Knife® participated in the study by delineating the target and performing the planning. The analysis of the delineated targets was based on the calculated 50% agreement volume, AV50, the encompassing volume and the common volume. The dosimetric implications were evaluated using the Conformity Index, Paddick Conformity Index and Gradient Index for each delineated target and the corresponding plan. The resulting high variability in target contouring showed in Paper I was not anticipated and a new study involving common SRS targets was initiated.

Paper II

The hypothesis that common targets would show a low disagreement in contouring variability was investigated in Paper II where six targets regarded as common, were chosen for analysis. The variability in the contouring was lower than for the complex targets but still much higher than expected. Another metric for comparing the targets based on the position of the Center of Mass, was used, and the results showed the highest disagreement in the Z-direction for the majority of cases.

Paper III

A similar analysis of the variability in delineation was performed for the OAR in SRS. The participants in this pilot study were intentionally given minimal instructions in terms of planning and contouring guidelines in order to generate results that would
reflect a clinical reality. The results showed a disagreement in structure contouring including several factors that were not expected and not shown in the analysis of targets. The availability of multiple choices of images for OAR contouring, the lack of clear specification regarding OAR tolerance doses, no indication on the part of the structure which should be contoured, all contributed to the very high variability in the structures contoured and led to formulating the need for OAR contouring guidelines.

Paper IV

This study introduced the radiobiological formalism for the evaluation of the treatment plans with respect to the probability of controlling tumours treated with SRS accounting for possible infiltrations of malignant cells beyond the margins of the delineated target. The study relates to the work in papers I and II as the high contouring variability for an invasive target such as anaplastic astrocytoma, where tumour cells could be present outside the tumour border defined on treatment planning MRI images, could dramatically impact on the calculated tumour control probability. The proposed framework could be used for the evaluation of stereotactic radiosurgery plans taking into account the possible infiltration of tumour cells around the visible target.

Paper V

With the implementation of hypo-fractionated treatment regimes in SRS and dedicated equipment and software, the impact of the fractionation scheme and hypoxic status on tumour control probability should be addressed in relation to the newly available treatment modalities. Three cases of metastases, of different sizes, were chosen for the study and the tumour control probability was calculated for different scenarios regarding the hypoxic status of cells and the fractionation scheme. It was shown that the response was worse with single fraction radiosurgery for large hypoxic tumours and that they could benefit from a fractionated treatment regime.
5. Outlook

Despite the long clinical experience and rapid development of stereotactic radiosurgery, little has been done to evaluate and overcome the uncertainty in target and Organs-at-Risk contouring. Several further studies, continuation of the current ones, are therefore planned. The Organs-at-Risk delineation variability study will be continued by expanding the current analysis. The new study will involve a number of experienced clinicians from well-established Gamma Knife centres from around the world who will have to perform the contouring and planning under strict instructions set by the working group on standardisation of OAR contouring for SRS.

Another sub-study that will be performed in the future will focus on the evaluation of the intra-observer variability where radiosurgery specialists will be evaluated with respect to their consistency in target contouring. This study is expected to emphasize the importance of training and experience in contouring.

Furthermore, the methods presented in this thesis for multicenter contouring variability are to be complemented by an analysis using the STAPLE method for both targets and Organs-at-Risk in the brain.

Finally, the ultimate aim of the planned future work is to develop together with the industrial collaborators from Elekta AB (Elekta Instrument AB, Stockholm, Sweden) an approach for robust treatment planning in Leksell GammaPlan® accounting for uncertainties in the target and OAR delineation in which the previously accumulated knowledge will be incorporated.
References


Kavanagh BD, McGarry RC and Timmerman RD. Extracranial radiosurgery (stereotactic body radiation therapy) for oligometastases. *Semin Radiat Oncol* 16:77-84 2006


Paddick I. and Lippitz B. A simple dose gradient measurement tool to complement the conformity index. *J Neurosurg*. 105 (Suppl.):S194-201 2006


Sayer FT., Sherman JH, Yen CP, Schlesinger DJ, Kersch R and Sheehan JP. Initial experience with the eXtend system: a relocatable frame system for multiple-session gamma knife radiosurgery. World Neurosurg. 75(5-6):665-72 2011


Toma-Dasu I, Sandström H, Barsoum P and Dasu A. To fractionate or not to fractionate? That is the question for the radiosurgery of hypoxic tumors. *J Neurosurg* 121 Suppl:110-5 2014


Weiss E and Hess CF. The impact of gross tumor volume (GTV) and clinical target volume (CTV) definition on the total accuracy in radiotherapy theoretical aspects and practical experiences. *Strahlenther onkol* 179(1):21-30 2003


Variability in target delineation for cavernous sinus meningioma and anaplastic astrocytoma in stereotactic radiosurgery with Leksell Gamma Knife Perfexion

Helena Sandström · Håkan Nordström · Jonas Johansson · Per Kjäll · Hidefumi Jokura · Iuliana Toma-Dasu

Received: 29 July 2014 / Accepted: 5 September 2014 / Published online: 23 September 2014 © Springer-Verlag Wien 2014

Abstract

Background Radiosurgery clinical practice relies on empirical observations and the experience of the practitioners involved in determining and delineating the target and therefore variability in target delineation might be expected for all the radiosurgery approaches, independent of the technique and the equipment used for delivering the treatment. The main aim of this study was to quantify the variability of target delineation for two radiosurgery targets expected to be difficult to delineate. The secondary aim was to investigate the dosimetric implications with respect to the plan conformity. The primary aim of the study has therefore a very general character, not being bound to one specific radiosurgery technique.

Materials and methods Twenty radiosurgery centers were asked to delineate one cavernous sinus meningioma and one astrocytoma and to plan the treatments for Leksell Gamma Knife Perfexion. The analysis of the delineated targets was based on the calculated 50 % agreement volume, $AV_{50}$. The $AV_{50}$ was compared to each delineated target volume by the concordance index and discordance index. The differences in location, size, and shape of the delineated targets were also analyzed using the encompassing volume compared to the common volume, i.e., the $AV_{100}$ of all delineated structures.

Results Target delineation led to major differences between the participating centers and therefore the $AV_{50}$ was small in comparison to each delineated target volume. For meningioma, the $AV_{50}$ was 5.90 cm$^3$, the $AV_{100}$ was 2.60 cm$^3$, and the encompassing volume was 13.14 cm$^3$. For astrocytoma, the $AV_{50}$ was 2.06 cm$^3$ while the $AV_{100}$ was extremely small, only 0.05 cm$^3$, and the encompassing volume was 43.27 cm$^3$. These variations translate into corresponding discrepancies in plan conformity.

Conclusions Significant differences in shape, size, and location between the targets included in this study were identified and therefore the clinical implications of these differences should be further investigated.

Keywords Leksell Gamma Knife Perfexion · Gamma Knife radiosurgery · Radiosurgery · Target delineation · Astrocytoma · Meningioma · Inter-observer variability

Introduction

The use of stereotactic radiosurgery (SRS) is continuously increasing due to the technical development which, among other things, has increased the conformity of the dose delivered, complemented by the positive clinical experience [1, 7, 9–13, 16, 17, 24]. The success of SRS depends on many clinical, dosimetric, and radiobiological factors. However, they may not all have the same impact on the outcome and therefore the key factors must be identified in order to develop strategies to maximize the therapeutic gain.

Delivering a highly conformal dose distribution to the target in one fraction or few fractions allowing, at the same time, the sparing of normal tissue and critical structures is part...
Table 1 Fractionation schemes and fundamental characteristics of stereotactic radiosurgery versus conventional radiation therapy

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Stereotactic radiosurgery</th>
<th>Conventional radiotherapy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of fractions</td>
<td>1</td>
<td>10–30</td>
</tr>
<tr>
<td>Dose per fraction (Gy)</td>
<td>20–40</td>
<td>1.8–3</td>
</tr>
<tr>
<td>Target definition and margins</td>
<td>GTV=CTV=PTV No margins</td>
<td>GTV&lt;CTV&lt;PTV Centimeters margins</td>
</tr>
</tbody>
</table>

of the basic concept of radiosurgery (RS) and SRS. Provided that the highly accurate RS and SRS equipment available today is used, planning and delivering the prescribed and highly conformal dose distribution is an achievable goal. However, without clinically accurate target definitions, the goal will not be achieved. As the planning target volume in radiosurgery is defined without the typical margins (Table 1) normally used in fractionated radiation therapy, the success of the RS/SRS critically depends on the accurate delineation of the target volume. In SRS, the uncertainties related to patient setup, target movement, and equipment are very small and, if applicable, the inter-fraction uncertainties are also very small, therefore no margin needs to be applied. However, in order to be able to explain the variations or disagreements in target delineation between various radiosurgery practitioners around the world, it could be discussed if not a de facto margin is used. This margin would clearly be related to the uncertainty in the perception of the target borders, which is especially critical for invasive tumors such as gliomas and other targets difficult to image, as well as the limitation imposed by normal tissue and critical structure sensitivity, similarly to surgery where a resected volume might differ according to the clinical observations during the surgical procedure.

The assessment of the quality of treatment planning for RS/SRS is usually performed with respect to the coverage, in terms of dose conformity, of the target together with minimization of dose to critical structures and organs in the proximity of the target. With an increasing ability to conform the dose distribution it becomes even more important to obtain a clinically accurate delineation of the target. In spite of its high clinical relevance, to the best of our knowledge there are only a few studies on the assessment of the variability in target delineation for radiosurgery [2, 19, 23], indicating that variation between observers in target delineation is to be expected and therefore is one of the current problems in clinical practice.

It is the aim of this study to assess the variability in target delineation for two types of tumors commonly treated with SRS using Leksell Gamma Knife Perfexion (LGK, Elekta Instrument AB, Stockholm, Sweden), namely anaplastic astrocytomas and meningiomas in the cavernous sinus region. Both of these tumors are prone to pose difficulties in delineation. Although this study is restricted to LGK, the assessment of the variability in target delineation has a general character, not being directly dependent on the technique used for delivering the single fraction of dose. The present study has also assessed the implications of the differences in the delineation on the conformity of the resulting plans performed for Leksell Gamma Knife Perfexion.

Materials and methods

A recurrent anaplastic astrocytoma and a cavernous sinus meningioma previously treated with Leksell Gamma Knife Perfexion at Furukawa Seiryo Hospital, Furukawa, Japan, were selected for this study. These cases are described below as case 1 and case 2, respectively.

Case 1 is a male patient, age 45, surgically treated for left thalamic anaplastic astrocytoma in 2004, followed by 72 Gy of accelerated hyperfractionated radiation therapy (1.2 Gy × 2/day, 5 days/week) and chemotherapy with Nimustine (ACNU). In 2009 the patient presented a recurrence in the contra-lateral thalamus in the head of the caudate nucleus as well as in the medial temporal lobe, which were treated with LGK. The prescribed dose was 16 Gy at the 50 % isodose. Several image sets, including MR (plain axial and gadolinium-enhanced axial+coronal T1-weighted images, axial T2-weighted images and fluid-attenuated inversion recovery-weighted images) and PET images (11C-methionine), were available for the target definition, delineation, and treatment planning for this patient. Examples of axial images of the astrocytoma case are shown in Fig. 1.

Case 2 is a female patient, age 75, with meningioma of the right cavernous sinus in the proximity to the right optical nerve. The patient presented with a dysfunction of cranial nerve III and VI causing diplopia and blepharoptosis, and was treated with LGK SRS in 2010 with 11 Gy to the 50 % isodose. MR (Gd-enhanced axial and coronal T1-weighted spin-echo images, coronal T2-weighted images) and CT (bone-window) images were available for this patient. Examples of axial and coronal images of the meningioma case are shown in Fig. 2.

Twenty centers chosen for their experience with LGK SRS participated in the study. Fourteen centers delineated the target for anaplastic astrocytoma and 16 centers delineated the target for cavernous sinus meningioma. The pre-planning images co-registered to the stereotactic images were sent to the participants. The participants were further instructed to perform the delineation and the planning according to the current practice at their respective clinic, with the obvious constraint that they could only use the provided image material. No hints...
or indications about the choice of the available images for delineation were given to the participants. Regarding the planning, the prescribed dose for case 1, was 16 Gy at the 50 % isodose and 11 Gy at the 50 % isodose for case 2. In order to avoid interpolation errors, the grid resolution for the target matrix was specified as 2.2 mm for case 1 and 1.3 mm for case 2, respectively, and the center point of the target matrices were fixed. There were no instructions specifying the constraints regarding organs at risk. Furthermore, no instructions were given regarding who should perform the target delineation or the treatment planning in order to mimic as close as possible the clinical routine at the various centers participating in the study and assess the variability in contouring the target as part of the regular radiosurgery work-flow. In addition, the participants were asked not to make a plan if they would refer the patient for another treatment modality at their clinic and not to radiosurgery. The resulting targets delineated by the 20 centers and the corresponding plans were collected and analyzed.

In order to determine the differences and similarities between the delineated targets, the planning images on which the targets were delineated by the participating centers were processed and binary 3D volumes were generated in MATLAB (MathWorks) by importing the DICOM structure-set files and

Fig. 1 Examples of the available images for case 1, anaplastic astrocytoma: a MRI-T1, b gadolinium-enhanced MRI-T1, c gadolinium-enhanced MRI-T1, d MRI-T2, e MRI-FLAIR, and f PET-methionine

Fig. 2 Examples of the available images for case 2, meningioma: a gadolinium enhanced axial MRI-T1, b CT, c gadolinium-enhanced coronal MRI-T1, d coronal MRI-T2
identifying the coordinates of the regions of interest. The resulting 3D volumes were, of course, identical to those from the treatment planning system. The voxels in these volumes were assigned a value of 1 if the voxel was located inside the planned target, whereas it was assigned a value of 0 if it was outside the planned target.

The encompassing volume, defined as the union of the volumes delineated by all participants was calculated. The 100 % agreement volume, AV100, defined as the intersection of all delineated volumes was also calculated. The definition of the volumes used is illustrated in Fig. 3. The large difference between the encompassing volume and AV100 is a result of a large variation in sizes, shapes, and positions, i.e., a large variability in the delineation of a target among the participants.

AV50 was calculated by adding all binary structures and defining the volume on which 50 % of the participants agreed. When 2 N delineations are performed, AV50 is the union of all voxels with values ≥ N. AV20 is in this work regarded as an average target (based on the assumption that all participants delineate with the same frame of reference).

The AV50 was compared to each individually delineated target by calculating the concordance and the discordance indices [4]. The concordance index (CCI) for each delineated target volume, V OBS, was calculated as (AV50 ∩ V OBS) / (AV50 ∪ V OBS). CCI is interpreted as the volume of intersection between AV50 and V OBS divided by the volume equal to the union of the two volumes. In the ideal case, there is no displacement of the two volumes and they are of the same size and shape and therefore CCI = 1. The discordance index (DCI) was determined as the fraction of the AV50, which is not covered by each individually delineated target volume,

\[ (AV50 - (AV50 \cap V_{OBS})) / AV50. \]

When the AV50 is completely covered by the individually delineated volume, DCI = 0.

In order to investigate the variability in the geometry and extension of the delineated targets, we propose a metric based on the agreement volume concept. To calculate this metric, all delineated volumes were expressed in the binary format explained above and added together. The 3D matrix thus constructed contains all AVi, where i ∈ [1, ..., number of participants], as well as the union of all volumes, the encompassing volume. Equivalent spherical volumes were then calculated for each AVi and therefore a radius R(AVi) could be assigned to each AVi. The largest difference in radius is then observed between R(AV100) and the encompassing volume, provided that all delineated volumes are not identical. Small differences between R(AVi) indicate that there is a corresponding small disagreement in the position, the shape, and the size of the target.

In order to assess the influence of the target delineation on the resulting treatment plan, one could use the parameter recommended by the Radiation Therapy Oncology Group (RTOG) [18] for the evaluation of stereotactic RT treatment plans, namely the conformity index (CI), which is calculated as the ratio between the volume of the target receiving a dose equal to or higher than the reference dose (prescription dose) and the target volume. Another measure of how well an RS plan conforms to the delineated target volume is the Paddick conformity index (PCI) [14, 20]. The PCI is calculated as \( \left( \frac{V_{TP}}{V_T} \right) \times \left( \frac{V_{TP}}{V_P} \right) \), were \( V_{TP} \) is the volume of target contained within the prescription isodose volume, \( V_T \) is the target volume, and \( V_P \) is the prescribed isodose volume. A measure of the dose fall-off outside the prescribed isodose volume is the gradient index (GI), which is calculated as the ratio between the prescribed isodose volume and the volume covered by half of the prescribed isodose [15]. All evaluation parameters determined for the targets delineated by participating centers are summarized in Table 2.

**Table 2** Plan evaluation parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Ideal value</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conformity index (CI)</td>
<td>( V_{TP}/V_T )</td>
<td>( V_{TP}=) Volume of target contained within prescription isodose volume, ( V_T=) target volume</td>
</tr>
<tr>
<td>Paddick Conformity Index (PCI)</td>
<td>( V_{TP}/V_T \times V_{TP}/V_P )</td>
<td>( V_{TP}=) Prescribed isodose volume, Measure of both under-treatment and over-treatment</td>
</tr>
<tr>
<td>Gradient Index (GI)</td>
<td>-</td>
<td>( V_{PRE}=) Prescribed isodose volume</td>
</tr>
<tr>
<td>Concordance index (CCI)</td>
<td>1</td>
<td>( AV_{50}=50 % ) agreement volume, ( V_{OBS}=) observed volume</td>
</tr>
<tr>
<td>Discordance index (DCI)</td>
<td>0</td>
<td>( AV_{50}-AV_{50} \cap V_{OBS} ) / AV50</td>
</tr>
</tbody>
</table>

**Fig. 3** The figure on the left displays the common volume, AV100, and the figure on the right shows the encompassing volume illustrated in one slice view.
Results

The differences and similarities in the delineated targets by all 20 participating centers in this study for the two analyzed cases are illustrated in Figs. 4, 5, and 6. The volumes of all delineated targets are given in Fig. 4 together with the AV$_{50}$. Figure 5 shows the overlap of the binary matrices collected from the participants for the astrocytoma case. White corresponds to the situation when all voxels are included in the target and black to all voxels outside the target, whereas shades of grey represent situations in between; in the slice shown, the highest level of agreement is 13 out of 14 participants. The corresponding illustration for the meningioma case is shown in Fig. 6 and the highest level of agreement in the slice displayed is 16. Note that in Figs. 5 and 6, both white and black correspond to full agreement; white when all agree that the voxel is inside the target, and black when all agree that the voxel is outside the target. Maximum disagreement is represented by the first step on the gray scale above black.

In Fig. 7, AV$_{MIN}$ to AV$_{MAX}$ are displayed for both cases, where AV$_{MIN}$ is equivalent to AV$_{100}$ normalized to AV$_{50}$ and AV$_{MAX}$ is equivalent to the encompassing volume normalized to AV$_{50}$. Figure 7 also displays the spherical representation of the agreement volumes, i.e., a graphic representation of the $R(AV)$ metric. The largest sphere in Fig. 7 represents the encompassing volume; the next smaller sphere represents the volume agreed by at least two participants, and so on. The lower panels of Fig. 7 illustrate a cross section of the spherical representation. Table 3 contains a comparison between all the volumes used in the analysis.

Graphical illustrations of the AV$_{50}$, encompassing and common volumes are shown in Figs. 8 and 9 as isosurface plots.

Large differences between AV$_{100}$ and encompassing volume were found for case 1, reflecting the higher complexity of this target, which also explains why a difference of 19.79 cm$^3$ between the largest and smallest delineated target was found. Three distinct outliers with respect to the volume were identified, but they were however not outliers with respect to position. The low median volume also indicates a higher number of delineated targets with small volumes for case 1. The main reason for why AV$_{50}$ is smaller than the median volume is then clearly the differences in positions and shapes of the individually delineated targets, and to a lesser extent the

---

Fig. 4 Bar plot of all volumes together with the AV$_{50}$ (marked in red) ordered by rising volume

Fig. 5 Example of the overlap between targets delineated by the individual participants in one slice for case 1, anaplastic astrocytoma. The grey scale shows the level of agreement. White corresponds to the highest level of agreement, which in this slide is 13 out of 14 participants, whereas the first step above black corresponds to total disagreement. Note that black (i.e., 0) also corresponds to total agreement

Fig. 6 Example of the overlap between targets delineated by the individual participants in one slice for case 2, meningioma. The grey scale shows the level of agreement. White corresponds to the highest level of agreement, which in this slide is 16, whereas the first step above black corresponds to total disagreement. Note that black (i.e., 0) also corresponds to total agreement
Fig. 7 The top figures display the percentage agreement volumes as a function of agreement volumes for a case 1 and b case 2. The middle and bottom figures illustrate the spherical representation of agreement volumes for c, e case 1, and d, f case 2. The bottom figures illustrate the cross section of the spherical volumes.

Differences in volumes. $AV_{100}$ is extremely small, only $\sim 0.1$% of the encompassing volume, as illustrated in Fig. 8 (bottom panel) and shown in Table 3. It is assumed that the main reason for this is the multi-focal appearance of this lesion.

Regarding the target delineations of the meningioma lesion and the resulting volumes, one observes that $AV_{50}$ is only slightly smaller than the mean and the median volume, as shown in Table 3. However, the difference between the $AV_{100}$ and the encompassing volume is 10.55 cm$^3$, i.e., almost two times $AV_{50}$; the interpretation of this result is that although the volumes are of the approximately the same size, the individually
delineated targets differ more in position than in shape and volume.

Statistics for $\text{AV}_{50}$ expressed in CCI and DCI, for both cases, is shown in Table 4. For case 1, two delineated targets resulted in a CCI value close to zero which, it is assumed, is due to the fairly high geometrical complexity of the astrocytoma lesion chosen for this study. Despite this, there were three delineated structures that gave a CCI value higher than 0.7, which means that those structures are similar to the $\text{AV}_{50}$. The DCI-statistics reveal that the average overlap between the delineated structures and $\text{AV}_{50}$ is around 70 % (of $\text{AV}_{50}$).

The analysis for case 2 shows that all of the delineated structures gave a CCI value higher than 0.5 and also that the mean and median CCI values are higher than 0.7. Higher values of the CCI index are the result of a better concordance between the delineated structures and $\text{AV}_{50}$, in shape, size, and position. DCI values also display a higher consistency among the delineated structures for the meningioma case, and here the average overlap between the delineated structures and $\text{AV}_{50}$ is close to 90 % of $\text{AV}_{50}$.

If it is assumed that the target has a unique shape, a unique volume, and a unique location, then it is reasonable to assume that the differences in the delineated targets we have quantified in this study translates into different qualities of the plans. Table 5 shows the values of the different conformity indices used for the two analyzed cases. The CI values for all meningioma plans are above 0.9, whereas they are consistently lower for the astrocytoma plan, again indicating the higher complexity of the astrocytoma target. PCI values as well as GI values show the same relationship; PCI (GI) values are consistently higher (lower) for the meningioma lesion than for the astrocytoma lesion.

### Table 3 Volumetric analysis of the delineated targets

<table>
<thead>
<tr>
<th></th>
<th>Anaplastic astrocytoma</th>
<th>Meningioma</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{AV}_{50}$ (cm$^3$)</td>
<td>2.06</td>
<td>5.90</td>
</tr>
<tr>
<td>Mean volume (cm$^3$)</td>
<td>5.90±6.21</td>
<td>6.53±0.74</td>
</tr>
<tr>
<td>Median volume (cm$^3$)</td>
<td>2.88</td>
<td>6.67</td>
</tr>
<tr>
<td>Min volume (cm$^3$)</td>
<td>1.66</td>
<td>4.96</td>
</tr>
<tr>
<td>Max volume (cm$^3$)</td>
<td>21.45</td>
<td>7.69</td>
</tr>
<tr>
<td>Common volume, $\text{AV}_{100}$ (cm$^3$)</td>
<td>0.05</td>
<td>2.60</td>
</tr>
<tr>
<td>Encompassing volume (cm$^3$)</td>
<td>43.27</td>
<td>13.14</td>
</tr>
</tbody>
</table>

Fig. 8 Case 1. Anaplastic astrocytoma: the top figure shows $\text{AV}_{50}$ (blue) versus encompassing volume (red), and the bottom figure shows the common volume, $\text{AV}_{100}$ (blue) versus encompassing volume (red).

Fig. 9 Case 2. Meningioma: the top figure shows the $\text{AV}_{50}$ (blue) versus encompassing volume (red) and the bottom figure shows the common volume, $\text{AV}_{100}$ (blue) versus encompassing volume (red).
For case 1, the highest GI value originates from a plan where a lower isodose is directed towards suspected tumor infiltration.

The analysis of the $R(AV_i)$ metric gives a difference in radius between $AV_{100}$ and the encompassing volume equal to 1.96 cm for case 1 and equal to 0.61 cm for case 2. All $R(AV_i)$ values for case 1 and 2 are displayed as bar plots in Fig. 10. The differences between two consecutive $R(AV_i)$ and the difference between case 1 and case 2 are clearly seen in these bar plots. It can be seen in Fig. 10 that all $R(AV_{i+1}) - R(AV_i)$ distances are less than 1 mm for case 2, while for case 1 there is a significant variation among these differences.

Discussion

Many studies have been published about variability in delineating the target for external fractionated radiotherapy [3, 5, 6, 8, 21–23]. However, only sparse information is available in the literature on the variations in target delineation for meningiomas and high-grade gliomas treated with SR/SRS, which is somewhat surprising when comparing with the number of publications dealing with the treatment of malignant glioma with LGK. It is important, however, to note that some of these articles do not clearly state the methodology of delineation of the target volume.

The degree of radicalism of the delineated target, ranging from small to very large targets, could have clinical consequences. Thus, a too small target could result in local recurrence of astrocytoma, for example. A very large delineated target might lead to normal or critical structure complications, but good control of the tumor, for example. Although his study was restricted to Gamma Knife Perfexion, the variability in the delineation of the target is independent of the radiosurgical technique and therefore a metric like the newly introduce $R(AV_i)$ should be of general interest.

Concerning the methodology of delineating the target, one important aspect is the use of automatic contouring tools and the profession or expertise behind the contours and plans. These issues were not investigated in this study, but it would be of high importance to distinguish if the outlying radically large delineated volumes are profession-dependent. A difficulty in incorporating this assessment in the current analysis is the fact that a resulting plan is often based on several professions in collaboration. To be able to isolate these factors, stricter instructions have to be sent to the participants, specifying who should perform the delineation. This might however result in lower-quality treatment plans, or in other words, in plans that would not be used in the clinical practice. The automatic delineation tools are often not completely trusted for highly complex targets, as in this study, but used more for assisting manual delineation.

The present study has pointed out the importance of stating the methodology of delineating target and the images used, especially for the treatment of invasive tumor such as glioma. A high inter-observer variation was noticed, as expected, due to the complexity of one of the targets involved and the intentionally relaxed set of instructions sent to the participants. Therefore, the participants in the study had the freedom to choose which images among those provided were to be used

<table>
<thead>
<tr>
<th>Table 4</th>
<th>Comparison of $AV_{50}$ to each individually delineated target by the concordance index and the discordance index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anaplastic astrocytoma</td>
<td>Meningioma</td>
</tr>
<tr>
<td></td>
<td>Concordance index (CCI)</td>
</tr>
<tr>
<td>Mean</td>
<td>0.45</td>
</tr>
<tr>
<td>Median</td>
<td>0.49</td>
</tr>
<tr>
<td>Max</td>
<td>0.85</td>
</tr>
<tr>
<td>Min</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 5</th>
<th>Plan conformity indices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anaplastic astrocytoma</td>
<td>Meningioma</td>
</tr>
<tr>
<td></td>
<td>Conformity index (CI)</td>
</tr>
<tr>
<td>Mean</td>
<td>0.91</td>
</tr>
<tr>
<td>Median</td>
<td>0.93</td>
</tr>
<tr>
<td>Max</td>
<td>0.96</td>
</tr>
<tr>
<td>Min</td>
<td>0.77</td>
</tr>
</tbody>
</table>
for the definition and delineation of the targets, as well as who within the clinical team should perform the delineation. This freedom on the participants’ side invariably propagates into questions concerning the results: was the possible or suspected infiltration region included or not in the assessments and delineations of the astrocytoma lesion? If so, what was the variation in size of this region among the different participants? Similar questions could be asked for the meningioma lesion, e.g., was the delineated target region affected by the proximity of the optic apparatus and did the presence of edema have an effect on how the participants viewed the target? The inherent uncertainties for these two targets have a large impact on the result and it should therefore be interpreted with care.

The aim of this study is to assess the variability in target delineation in the absence of a very strict set of instructions and to develop a methodology for the analysis of the variability in target delineation. The questions posed above are therefore left without answers in this study and therefore a continuation of the study is planned. Furthermore, future plans include the assessment of the influence of the accuracy in delineating the target beyond the volumetric and dosimetric analysis performed in this study by also performing a radiobiological evaluation with respect to the expected control of the tumor and normal tissue damage. We see merits with the introduced \( R(\text{AV}_i) \) metric, despite its complexity, since it offers an easy-to-handle tool for the assessment of interobserver target variability, which might have the potential to help choosing the margin for radiosurgical targets if proven necessary, and therefore we will work on developing this formalism further.

Conclusions

The results of this study emphasize the importance of understanding target delineation in stereotactic radiosurgery and particularly in SRS performed with very high accuracy devices, such as the Leksell Gamma Knife Perfexion. Differences in target delineation were studied with different techniques, among which one is novel, the equivalent radius metric. Central to the analysis of the delineated targets was the calculated 50 % agreement volume structure, \( \text{AV}_{50} \), which is assumed to constitute the “true” target if all participants delineate with the same clinical information, experience, and clinical endpoint. Very general instructions were given to the participants with the purpose of minimizing the influence on their clinical work, and the conclusion is that the differences
between the targets delineated for the investigated cases to a high degree of likelihood are clinically significant. We base this statement on the significant variability we found in the differences in location, shape, and size of the delineated targets, which naturally have a direct effect on values of the conformity indices used in the study.

Acknowledgments

Financial support from the Cancer Research Funds of Radiumhemmet is gratefully acknowledged.

Conflict of interest

None.

Informed consent

Informed consent was obtained from all individual participants included in the study.

References


Comments

In this potentially very important paper, the authors compared the outlines of two tumors targeted with stereotactic radiosurgery as performed in 20
radiosurgery units and demonstrated that there was significant disagree-
ment between the delineated volumes and thus the eventual treatment
plan. As all departments had the same radiological material available, the
differences resulted from different levels of experience (of the depart-
ments or at least the individuals), personality of the practitioner, desired
radicality, acceptability of complications, etc. As the authors acknowl-
dedge, even with the same equipment, different practitioners would deliver
different treatments for the same pathology. This is of course analogous to
the differences of how the same tumor would be operated on by different
surgeons. Furthermore, the difference starts already at the stage when the
radiological material is analyzed, even more in the case of radiosurgery
than in the case of open surgery where direct vision is also applied during
the procedure. Just as the different neurosurgeons reading this paper, the
practitioners in this study will have different personality profiles. As the
authors implied, a radiation oncologist may be more willing to use a wider
margin than a neurosurgeon and even in a multidisciplinary team such
differences do not disappear. Even amongst neurosurgeons, some are
more radical than others. Ideally, the radiology should be interpreted by
a radiologist and radicality should only appear at this stage of choosing a
radiation dose, but there is a temptation not even to mark as pathology
something that one does not intend to treat. It is almost arbitrary how far,
for instance, one plans to follow the dural tail of a meningioma. Naturally,
these considerations would not arise in a well-defined pathology, e.g.,
vestibular schwannoma. Also, even a Gamma knife treatment with the
sharpest dose fall-off has a small marginal effect beyond the prescription.
The results presented in this paper could be taken as quite shocking,
particularly for the anaplastic astrocytoma where some delineated 13
times as much volume as the most conservative team. However, one also
has to be aware that the next step, deciding the dose, may dramatically
diminish this difference, because the team outlining a large volume with a
margin around the tumor may prescribe a very low dose and vice versa.
The very interesting new indices the authors introduced, CCI and DCI,
would not take this into consideration, and perhaps a further index
incorporating a treatment plan may be developed. At the end, all that
matters is the final clinical outcome and of course the same patient will
not be treated by 20 different departments. One has to acknowledge that
radiosurgery, just as open surgery, is not an absolute exact science, but an
art, and in spite of technological advances, careful judgment based on
experience will remain the determinant of what we achieve.

Andras Kemeny,
Sheffield, UK

The authors present an interesting and important study on target
contouring in radiosurgery. They asked several centers to contour and
do the treatment planning for two identical targets. The various centers
received identical imaging and the required prescription doses were the
same for all centers. The authors then compared the results. Unfortunat-
ely, they chose two cases that were not really typical, considering everyday
radiosurgical practice. The cavernous sinus meningioma is rather large,
receiving a comparatively low dose and the anaplastic astrocytoma is
rather small, receiving a comparatively high dose. In fact, the mean
volume of the cavernous sinus meningioma is smaller than that of the
anaplastic astrocytoma. In addition, the authors chose with the cavernous
sinus meningioma, a case which is best contoured in a coronal plane
and with the anaplastic astrocytoma, a case which may as well be contoured
in an axial plane. The input of an experienced neurosurgeon when selecting
the exemplary cases would have been of great advantage. However,
setting aside these shortcomings, the study's results are interesting and
important enough to justify their publication. The differences in target
volumes are surprisingly large, with a standard deviation of ±0.74 of the
mean volume in the meningioma case and ±6.21 in the astrocytoma case.
Not surprisingly, the differences are much smaller in the meningioma case
where there is less room for ambiguity concerning target definition. What
can we learn from this study's results? (1) Target contouring in radiosur-
gery needs our attention. Contouring should be done by a specialist with
the necessary knowledge of anatomical and disease-specific details. Only
a neurosurgeon or a neuroradiologist has that knowledge. (2) Contouring
needs to be clearly separated from treatment planning. The target should
be delineated uncompromised by possible treatment plans. (3) Some
targets are better contoured in coronal than in axial planes. (4) Imaging
modalities are important for contouring. In my view, FLAIR sequences
are crucial in target contouring of gliomas. (5) Not all the differences in
target volumes are bad. Some may just reveal different treatment strate-
gies. Some radiosurgeons may prefer to treat gliomas according to C.
Duma’s “Leading Edge” principles while others may not. None of the
resulting treatment plans are wrong according to today’s standards of care.
(6) This study exemplifies the importance of contouring guidelines for
target volumes according to the contouring guidelines for organs at risk
(OAR). OAR guidelines are still missing, but they are worked out by the
international OAR group, which met in Toronto in 2013 and in New York
City in 2014. A target-contouring group still needs to be launched.

Thomas Mindermann,
Zurich, Switzerland
A Multi-Institutional Study of the Variability in Target Delineation for Six Targets Commonly Treated with Radiosurgery

Helena Sandström1, Hidefumi Jokura2, Iuliana Toma-Dasu1

1Medical Radiation Physics, Stockholm University and Karolinska Institutet, Stockholm, Sweden
2Suzuki Memorial Gamma House, Furukawa Seiryo Hospital, Osaki, Japan

ABSTRACT

The aim of the study was to quantify the variability in target delineation for six targets regarded as common in SRS, one cavernous sinus meningioma, one pituitary adenoma, one vestibular schwannoma and three cases of metastases under the hypothesis that common targets would show a low disagreement in contouring variability in comparison to complex targets that were previously investigated. The analysis of target delineation variability is based on agreement volumes derived from overlapping structures following a previously developed method. The position of the Center of Mass was also determined and compared for the delineated targets. The results showed that the variability in the contouring was indeed lower than for the complex targets but still much higher than expected and therefore further work in standardizing the contouring practice in SRS is warranted with the aim of reducing the inter-observer contouring variability.
INTRODUCTION

Stereotactic radiosurgery (SRS) is a treatment technique known for its high capability of delivering highly conformal dose to the target at the same time allowing normal tissue sparing due to the steep dose-fall off outside the prescribed dose volume. However, it does not explicitly take into account the accuracy in target contouring and therefore the outcome of the treatment is highly dependent on the local clinical practice. Because of the high dose fall-off outside the prescribed isodose, any inaccuracy in target contouring could result in a potential lower probability of tumour control or increased normal tissue toxicity. A recent study by Sandström et al. (2014) showed indeed a very high variability in target contouring for two complex lesions in the brain, one case of cavernous sinus meningioma and one case of anaplastic astrocytoma. Twenty centers using Leksell Gamma Knife® for performing SRS provided contours and treatment plans for the two cases and large differences were found within these contours with respect to the size, location and shape of the target. These results indicate a possible impact on the treatment outcome. Other studies have as well investigated the variety of intracranial target contouring differences (Heesters et al. 1993, Leunens et al. 1993, Weltens et al. 2001, Buis et al. 2005, Stanley et al. 2013). Methods of comparing multi-observer contouring data often rely on the analysis of overlapping structures (Rasch et al. 1997, Logue et al. 1998, Yamamoto et al. 1999, Fox et al. 2005, Struikmans et al. 2005, Song et al. 2006, Voroney et al. 2006, Landis et al. 2007, Petersen et al. 2007, Hurkmans et al. 2009, Li et al. 2009, Vorwerk et al. 2009, Louie et al. 2010, Rasch et al. 2010, Lütgendorf-Caucig et al. 2011, Altorjai et al. 2012, Sandström et al. 2014, Sandström et al. 2015). These methods depend on the number of contours/segmentations analysed and therefore set the limit on the possibility of inter-study comparison.

The purpose of this study is to evaluate multi-center contouring data for six targets in the brain commonly treated with SRS using the methods and the metrics proposed by Sandström et al. (2014) together with the evaluation of the position of the Center of Mass (COM).

MATERIAL AND METHODS

Clinical cases and imaging

Patient data consists of a cavernous sinus meningioma (case 1), pituitary adenoma (case 2), vestibular schwannoma (case 3) and three cases of metastases. They were previously treated at Suzuki Memorial Gamma House at Furukawa Seiryo Hospital in Osaka Japan.

Case 1, cavernous sinus meningioma, was initially diagnosed and treated in 2012. The following image data sets were available, also described in the study by Sandström et al. (2015): axial contrast-enhanced T1-weighted MR images (slice thickness 1 mm), axial T2-weighted MR images fused with CT (slice thickness 1 mm), fused axial image of contrast enhanced MRI T1-weighted with CT (slice thickness 1 mm), axial MRI T1-weighted image (slice thickness 1 mm), coronal MRI image (slice thickness 1 mm) and CT (slice thickness 0.6 mm). Examples of the images available for case 1 are shown in Figure 1.

Case 2, non-functioning pituitary adenoma, was diagnosed in 2002 followed by subtotal removal in 2004 and treated with SRS in 2013. Images available for contouring are: axial contrast enhanced MRI T1-weighted image (slice thickness 2.2 mm), fused axial image of contrast enhanced MRI T1-weighted with CT (slice thickness 2.2 mm), axial MRI T1-weighted image (slice thickness 2.2 mm), fused coronal image of contrast enhanced MRI T1-weighted with CT (slice thickness 2.2 mm), coronal MRI T1-weighted image (slice thickness 2.2 mm), fused coronal image of contrast enhanced MRI T1-weighted with CT (slice thickness 2.2 mm), coronal contrast enhanced MRI T1-weighted image (slice thickness 2.2 mm), coronal MRI T2-weighted image (slice thickness 2.2 mm), coronal MRI T1-weighted image (slice thickness 2.2 mm) and CT (slice thickness 0.6 mm). Examples of the images available for case 2 are shown in Figure 2.
Figure 1. Example images from Leksell GammaPlan® for Case 1: (A) T1-weighted MRI, (B) contrast enhanced T1-weighted MRI, (C) T1-weighted MRI fused with CT image, (D) T2-weighted MRI fused with CT, (E) coronal MRI and (F) CT image.

Figure 2. Example images from Leksell GammaPlan® for Case 2: (A) T1-weighted MRI, (B) contrast enhanced T1-weighted MRI, (C) T1-weighted MRI fused with CT image, (D) CT image, (E) coronal contrast enhanced T1-weighted MRI, (F) coronal T2-weighted MRI and (G) coronal contrast enhanced T1-weighted MRI fused with CT.
Case 3, a vestibular schwannoma, was diagnosed and treated in 2013. Available images for contouring were: axial MRI T1-weighted image (slice thickness 1 mm), fused axial image of MRI T2-weighted with CT (slice thickness 1 mm), fused axial image of contrast enhanced MRI T1-weighted with CT (slice thickness 1 mm), coronal MRI T2-weighted image (slice thickness 1 mm), axial MRI T2-weighted image (slice thickness 1 mm), axial contrast enhanced MRI T1-weighted image (slice thickness 1 mm) and CT (slice thickness 0.6 mm). Examples of the images available for case 3 are shown in Figure 3.

Case 4, a metastasis which in this study is classified as large (approximate diameter around 30 mm), was diagnosed and treated in 2008. Available images for contouring are: axial MRI T1-weighted image, axial contrast enhanced MRI T1-weighted image, Fluid Attenuated Inversion Recovery (FLAIR) MRI image and coronal MRI image. Slice thicknesses are 7 mm, 2 mm, 4 mm and 2 mm respectively. Examples of the images available for case 4 are shown in Figure 4.

Case 5, a metastasis which in this study is classified as medium (approximate diameter around 20 mm), was diagnosed and treated in 2012. Available images for contouring are: axial MRI T1-weighted image, axial contrast enhanced MRI T1-weighted image, Fluid Attenuated Inversion Recovery (FLAIR) MRI image and coronal MRI image. Slice thickness is 4 mm for the FLAIR sequence and 2 mm for the other images. Examples of the images available for case 5 are shown in Figure 5.
Case 6, a metastasis which in this study is classified as small (approximate diameter around 13 mm), was diagnosed and treated in 2013. Available images for contouring are: axial MRI T1-weighted image (slice thickness 2 mm), axial contrast enhanced MRI T1-weighted image (slice thickness 2 mm), axial Fluid Attenuated Inversion Recovery (FLAIR) MRI image (slice thickness 4 mm) and coronal MRI image.

Examples of the images available for case 6 are shown in Figure 6.

**Data collection**

All images for one particular case were co-registered and anonymized before distribution to twelve centers using Leksell Gamma Knife® for performing SRS. The data was sent as LGP-files, easily imported to the treatment planning system (Leksell GammaPlan®, Elekta Instrument AB, Stockholm, Sweden). The instructions given to the participants were to contour the target and OARs according to the local clinical practice and to perform at least one treatment plan for each case. The skull contour was pre-specified before distribution of data.

Information about the participants was collected as described by Sandström et al. (2015) and included profession of planner, experience, planning methodology and general comments on the cases provided.

**Data analysis**

The methodology used for analysis was previously described by Sandström et al. (2014) and Sandström et al. (2015). The contours were exported from Leksell GammaPlan® (LGP) as Digital Imaging and Communications in Medicine (DICOM)-objects and processed in MATLAB (Mathworks, Inc).

Figure 7 illustrates examples of contoured structures (left panel) and the variability in target outlining in one slice (right panel).

The 50% agreement volume, $AV_{50}$, the structure that 50% of participants agree on, was calculated for each of the 6 cases as described in Sandström et al. (2014). The contours were further analysed and the encompassing volume, $AV_{100}$, of all delineated structures was determined. Figure 8 shows the binary representation of the variability in one slice where black and white indicate complete or highest level of agreement in one slice. Levels of grey illustrate the different levels of agreement from $N_i$ (complete) to $N_i^{-1}$ (full disagreement) where $N$ is number of contours for target i. $N_i$ is related to $AV_{100}$ and $N_i^{-1}$ to $AV_{100}/N_i$, the intersect and the union of the contoured structures, respectively.
Figure 7. Examples of contoured structures (left panel) and overlapping contours in one slice (right panel) for (A,B) cavernous sinus meningioma, (C,D) pituitary adenoma, (E,F) vestibular schwannoma, (G,H) metastasis (L), (I,J) metastasis (M), (K,L) metastasis (S).
Figure 8. One slice of the binary agreement matrix for (A) cavernous sinus meningioma, (B) pituitary adenoma, (C) vestibular schwannoma, (D) metastasis (L), (E) metastasis (M1), (F) metastasis (S).

The ratio of volume \( N_i \) with \( N_i^{-1} \) is the Agreement Volume Index (AVI) which is a non-negative value between 0 and 1 and gives the level of agreement with respect to the volume, size and shape of contoured structures, previously described by Sandström et al. (2015).

The Center of Mass (COM) is the weighted center point of a structure. The position of COM, a 3D coordinate, is calculated for each individual contoured structure. From these coordinates, the largest difference in X, Y or Z-direction is calculated and given in mm.

The treatment plan evaluation parameters, target coverage, selectivity and gradient index were also assessed for each of the targets which are parameters given in the TPS. The treatment time was also calculated for each plan based on the same dose-rate, 2.478 Gy/min (corresponding to replacement of sources in January 2013 with an initial dose-rate of 3.578 Gy/min at time of replacement).

RESULTS

Twelve Gamma Knife centers participated in the study (from Greece, Norway, Czech Republic, Japan, South Korea, Canada, United Kingdom and USA). 32 professionals were involved including 13 neurosurgeons (5-21 years of experience), 12 physicists (3-20 years of experience), 3 radiation oncologists (6.5-24 years of experience), 3 radiologists (7-11 years of experience) and 1 neuroradiologist (10 years of experience). A total of 71 treatment plans with contours were distributed back to be a part of the study; 12 plans for case 1, 13 plans for case 2, 12 plans for case 3, 10 plans for case 4, 12 plans for case 5 and 12 plans for case 6.

Image data sets used for contouring

A summary of the images used for contouring target are shown in Table 1

For case 1 the majority of participants contoured the cavernous sinus meningioma based on the contrast enhanced T1-weighted MRI or the fused image of CT with contrast enhanced T1-weighted MRI. One structure was contoured on the fused image of CT with T2-weighted MRI. For case 2, pituitary adenoma, the contours were drawn on T1-weighted images, contrast enhanced or non-contrast enhanced, coronal or axial together with the use of a fused image of CT with contrast enhanced T1-weighted MRI. The agreement is high for cases 3-6, where the majority or all contours are drawn on axial contrast enhanced T1-weighted MRI.
<table>
<thead>
<tr>
<th></th>
<th>Contrast enhanced axial T1-weighted MRI</th>
<th>Fused image of contrast enhanced axial T1-weighted MRI with CT</th>
<th>Fused image of axial T2-weighted MRI with CT</th>
<th>Contrast enhanced coronal T1-weighted MRI</th>
<th>Axial T1-weighted MRI</th>
<th>Fused image of coronal contrast enhanced T1-weighted MRI with CT</th>
<th>Coronal T1-weighted MRI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cavernous sinus meningioma</td>
<td>6</td>
<td>5</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Pituitary adenoma</td>
<td>3</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Vestibular schwannoma</td>
<td>10</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Metastasis (L)</td>
<td>10</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Metastasis (M)</td>
<td>12</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Metastasis (S)</td>
<td>11</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Figure 9. Range of volumes contoured together with the AV50 for (A) cavernous sinus meningioma, (B) pituitary adenoma, (C) vestibular schwannoma, (D) metastasis (L), (E) metastasis (M), (F) metastasis (S).
Figure 10. Illustration of the AV$_{100}$ (blue) and AV$_{50}$ as 3-dimensional surfaces together with the Center of Mass (red points) for (A) cavernous sinus meningioma, (B) pituitary adenoma, (C) vestibular schwannoma, (D) metastasis (L), (E) metastasis (M), (F) metastases (S).

**Analysis of the contouring variability**

The range of contoured volumes together with respective AV$_{50}$ is shown in Figure 9. The range of volumes is higher for case 1 (5.29-7.80 cm$^3$), case 4 (10.30-14.55 cm$^3$) and case 5 (1.27-3.33 cm$^3$) while the range of volumes contoured for case 2, case 3 and case 6 are lower: 1.67-2.15 cm$^3$, 3.53-4.48 cm$^3$ and 1.42-2.26 cm$^3$ respectively. The AV$_{50}$ is lowest for the large metastasis (case 4), only 3.90 cm$^3$ indicating a high variability in the structure itself with regard to position and shape. All other cases are
Table 2. Summary of the results including the volumes contoured, $AV_{50}$, $AV_{100/N}$, $AV_{100}$, the agreement volume index $AV_{100}/AV_{100/N}$ and treatment planning parameters: range of treatment time, prescribed dose, coverage, selectivity and gradient index.

<table>
<thead>
<tr>
<th></th>
<th>Range of contoured volumes (cm$^3$)</th>
<th>$AV_{50}$ (cm$^3$)</th>
<th>$AV_{100/N}$ (cm$^3$)</th>
<th>$AV_{100}$ (cm$^3$)</th>
<th>$AV_{100}/AV_{100/N}$ (cm$^3$)</th>
<th>Treatment time (min) (median)</th>
<th>Prescribed dose (Gy) (median)</th>
<th>Coverage (median)</th>
<th>Selectivity (median)</th>
<th>Gradient index (median)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cavernous sinus meningioma</td>
<td>5.29-7.80</td>
<td>5.52</td>
<td>11.06</td>
<td>2.48</td>
<td>0.22</td>
<td>77.5-319.6 (114.2)</td>
<td>12-16 (13.25)</td>
<td>0.95-0.99</td>
<td>0.73-0.90</td>
<td>2.58-2.74 (2.61)</td>
</tr>
<tr>
<td>Pituitary adenoma</td>
<td>1.67-2.15</td>
<td>1.80</td>
<td>2.98</td>
<td>1.09</td>
<td>0.37</td>
<td>43.6-115.0 (63.8)</td>
<td>12-18 (15)</td>
<td>0.89-0.99</td>
<td>0.75-0.94</td>
<td>2.53-2.89 (2.68)</td>
</tr>
<tr>
<td>Vestibular schwannoma</td>
<td>3.56-4.48</td>
<td>3.90</td>
<td>5.40</td>
<td>2.68</td>
<td>0.50</td>
<td>49.8-136.8 (81.9)</td>
<td>11-13 (12.25)</td>
<td>0.94-1.00</td>
<td>0.79-0.97</td>
<td>2.62-3.06 (2.71)</td>
</tr>
<tr>
<td>Metastasis (L)</td>
<td>10.30-14.55</td>
<td>3.90</td>
<td>15.32</td>
<td>9.46</td>
<td>0.62</td>
<td>39.3-120.6 (72.7)</td>
<td>15-22 (18)</td>
<td>0.95-1.00</td>
<td>0.80-0.94</td>
<td>2.62-3.04 (2.65)</td>
</tr>
<tr>
<td>Metastasis (M)</td>
<td>1.27-3.33</td>
<td>2.94</td>
<td>4.29</td>
<td>0.68</td>
<td>0.16</td>
<td>23.9-72.9 (57.8)</td>
<td>18-25 (21.5)</td>
<td>0.92-1.00</td>
<td>0.50-0.94</td>
<td>2.45-2.98 (2.54)</td>
</tr>
<tr>
<td>Metastasis (S)</td>
<td>1.42-2.26</td>
<td>1.51</td>
<td>2.75</td>
<td>1.01</td>
<td>0.37</td>
<td>35.0-119.8 (60.1)</td>
<td>18.5-25 (22)</td>
<td>0.97-1.00</td>
<td>0.50-0.92</td>
<td>2.46-3.61 (2.72)</td>
</tr>
</tbody>
</table>
presented with an AV$_{50}$ representative of the size of contoured volumes. These values are shown in Table 2 for all six cases.

Figure 10 shows 3D illustrations of the AV$_{100}$ (blue) together with AV$_{100/N}$ (green). The position of COM is as well shown in Figure 10.

The largest deviation, for the COM, in any direction is shown in Table 3. The largest deviation between contoured structures was observed in Z-direction (cranio-caudal) for case 1, case 2, case 4 and case 6 while case 3 and case 5 have the largest deviation in the X-direction (anterior-posterior).

**Discussion**


Stanley et al. (2013) observed a high contouring variability for fourteen metastases delineated by 8 physicians for stereotactic radiosurgery and suggested that methods should be adopted to reduce variations in tumour definition. This suggestion was further supported by Hoang Duc et al. (2015) in a study where an automatic segmentation tool was evaluated for brain and head-and-neck OAR. Clinically accepted segmentations of the relevant OARs were obtained, including the brainstem and optic chiasm. Njeh (2008) also concluded that tumour definition is one source of uncertainty in radiotherapy and highlights the use of multiple imaging modalities, such as co-registration of CT with MRI or functional imaging such as positron emission tomography (PET), together with education and collaboration across specialties as possible factors which can reduce the contouring variability.

With no margins applied to a contoured target, as in SRS, the differences in contouring practices have a clinical impact in terms of reduced tumour control probability and normal tissue toxicity. Furthermore, as SRS techniques transition towards fractionated treatment regimes, accuracy in target contouring is especially important to lower the contouring inconsistency between fractions.

**Conclusions**

Six common SRS targets were analysed with respect to the contouring variability. The results showed considerable differences in the size, shape and position of the delineated targets highlighting the importance of guidelines for target contouring in SRS, as well as training and experience for professionals working in SRS.
REFERENCES


institutions around the world. In review Radiotherapy and Oncology 2015


Yamamoto M, Nagata Y, Okajima K et al. Differences in target outline delineation from CT scans of brain tumours using different methods and different observers. Radiother Oncol 50(2):151-156 1999


Preliminary Assessment of Organs-at-Risk Contouring Practices in Radiosurgery Institutions around the World

Helena Sandström MSc¹, Caroline Chung MD FRCPC CIP², Hidefumi Jokura MD³, Michael Torrens ChM FRCs⁴, David Jaffray PhD⁵, Iuliana Toma-Dasu PhD¹

¹Medical Radiation Physics, Stockholm University and Karolinska Institutet, Stockholm, Sweden
²Department of Radiation Oncology, University of Toronto and University Health Network-Princess Margaret Cancer Centre, Toronto, Canada
³Suzuki Memorial Gamma House, Furukawa Seiryo Hospital, Osaki, Japan
⁴Gamma Knife Department, Hygeia Hospital, Athens, Greece
⁵Department of Radiation Oncology, Medical Biophysics, IBBME, University of Toronto and University Health Network-Princess Margaret Cancer Centre/TECHNA, Toronto, Canada

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.

Material and Methods

Imaging datasets for typical lesions (cavernous sinus meningioma, vestibular schwannoma, pituitary adenoma) treated with Leksell Gamma Knife Perfexion were circulated to 12 centres. Observers were asked to contour the target and OARs as per their standard clinical practice. The analyzed parameters were the intersection (AV₁₀₀), union volumes (AV₁₀₀/N) and the 50% Agreement volume (AV₅₀). The ratio of AV₁₀₀ and AV₁₀₀/N (the Agreement Volume Index, AVI) was used as a measure of agreement level. 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 AV₁₀₀ 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.

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.
INTRODUCTION

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 [3,4,5,6,7,8]. Sandström et al. [8] investigated the variability in target volume contouring for one case of cavernous sinus meningioma and 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 efforts to standardize OAR contouring for radiotherapy planning of extracranial sites including head and neck, breast, gynecological, prostate and anal cancers [9,10,11]. To the best of our knowledge, this is the first study to evaluate the variability in practice of OAR contouring for intracranial radiosurgery.

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 evaluate 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 T1-weighted MR-images, axial T2-weighted MR-images fused with CT, fused axial image of contrast-enhanced MRI T1-weighted with CT, axial MRI T1-weighted image, axial MRI T2-weighted image, coronal MRI-image and CT. All images except for the CT image have a slice thickness of 1 mm except for the CT image 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 T1-weighted image, fused axial image of contrast-enhanced MRI T1-weighted with CT, axial MRI T1-weighted image, fused coronal image of contrast-enhanced MRI T1-weighted with CT, coronal contrast-enhanced MRI T1-weighted image, coronal MRI T2-weighted image, coronal MRI T1-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 T1-weighted image, fused axial image of MRI T2-weighted with CT, fused axial image of contrast-enhanced MRI T1-weighted with CT, coronal MRI T1-weighted image, axial MRI T2-weighted image, axial contrast-enhanced MRI T1-weighted image and CT. The slice thicknesses of these data sets are 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.

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 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 Figure 1 where white and black correspond to complete or highest level of agreement in that slice and grey corresponds to partial agreement. The 50%-agreement volume, AV50, 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 AV100 and AV100/N could be extracted. These represent the common volume, the intersection, 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.

The dosimetric analysis was performed with respect to the maximum dose to the structures from the corresponding individual dose-plan, the maximum dose to AV50 and the maximum dose to AV100/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.

RESULTS

Twelve Gamma Knife centers from Greece, Norway, Czech Republic, Japan, South Korea, Canada, United Kingdom (2) and USA (4) participated in this study. A total of 13 neurosurgeons, 12 physicists, 3 radiation oncologists, 1 neuroradiologist and 3 radiologists participated from these centers. Experience ranged across participants: participating neurosurgeons reported 5 to 21 years of experience, physicists 3 to 20 years, radiation oncologists 6.5 to 24 years, radiologists 7 to 11 years and the neuroradiologist had 10 years of experience.

Variability in OAR Nomenclature and the OARs Contoured as Part of Clinical Practice

The analyzed OARs together with number and percentage of participants contouring each structure are summarized in Table 1.
Table 1. Contouring and dosimetric results

<table>
<thead>
<tr>
<th></th>
<th>Cavernous sinus meningioma</th>
<th>Pituitary adenoma</th>
<th>Vestibular schwannoma</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Left optic nerve</td>
<td>Left optic tract</td>
<td>Right optic nerve</td>
</tr>
<tr>
<td>Volumes (cm³)</td>
<td>0.33-0.61</td>
<td>0.06-0.21</td>
<td>0.31-0.48</td>
</tr>
<tr>
<td>AV₅₀ (cm³)</td>
<td>0.46</td>
<td>0.11</td>
<td>0.44</td>
</tr>
<tr>
<td>AV₁₀₀ (cm³)</td>
<td>0.03</td>
<td>0.0</td>
<td>0.12</td>
</tr>
<tr>
<td>AV₁₀₀/N (cm³)</td>
<td>1.11</td>
<td>0.25</td>
<td>0.92</td>
</tr>
<tr>
<td>AVI</td>
<td>0.029</td>
<td>0.0</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>The range of the highest dose to each individual structure (Gy)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Highest dose to encompassing volume based on each plan (Gy)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Number of contoured structures (percentage)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nomenclature</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Nomenclature**

- Opt: Optic nerve
- LT-ON: Left optic nerve
- nll: Nontumoral left lobe
- Chl: Cochlea
- T2: Trigeminal nerve root
- Lt: Left
- o1: Oculomotor nerve
- lon: Optic nerve
- oar: Oculomotor nerve
- oc: Optic chiasm
- lot: Optic nerve
- ron: Optic nerve
- Lt ON: Left optic nerve
- Rt ON: Right optic nerve
- vispath: Optic chiasm
- Opt ax: Optic tract
- Nontumoral right lobe
- o1: Oculomotor nerve
- lon: Optic nerve
- oar: Oculomotor nerve
- oc: Optic chiasm
- lot: Optic nerve
- ron: Optic nerve
- Lt ON: Left optic nerve
- Rt ON: Right optic nerve
- Opt: Optic nerve
- opt ax: Optic tract
- Lt ON: Left optic nerve
- opt: Optic nerve
- opt ax: Optic tract
- nll: Nontumoral left lobe
- Lt ON: Left optic nerve
- o1: Oculomotor nerve
- lon: Optic nerve
- oar: Oculomotor nerve
- oc: Optic chiasm
- lot: Optic nerve
- ron: Optic nerve
- Lt ON: Left optic nerve
- Rt ON: Right optic nerve
- opt: Optic nerve
- opt ax: Optic tract
- AV: Average volume
- AVI: Average interval volume
- AV₅₀: Average volume based on the 50th percentile
- AV₁₀₀: Average volume based on the 100th percentile
- AV₁₀₀/N: Average volume based on the 100th percentile normalized to the encompassing volume
- G: Gamma value
- Gy: Gray equivalent dose
Table 2. Summary of the imaging datasets used for OAR delineation. The image sets used by the majority of participants for each OAR are marked with *.

<table>
<thead>
<tr>
<th>Imaging datasets utilized for OAR contouring</th>
<th>T1+C (CT+T2)</th>
<th>FA2 (CT+T2)</th>
<th>FA (CT+T1+C)</th>
<th>T1</th>
<th>T2</th>
<th>COR</th>
<th>CT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cavernous sinus meningioma</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left optic nerve</td>
<td>4*</td>
<td>3*</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Right optic nerve</td>
<td>2*</td>
<td>2*</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Chiasm</td>
<td>3*</td>
<td>3*</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Left optic tract</td>
<td>2*</td>
<td>3*</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Right optic tract</td>
<td>2*</td>
<td>2*</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td><strong>Vestibular schwannoma</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cochlea</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>3*</td>
<td>-</td>
<td>2*</td>
</tr>
<tr>
<td><strong>Pituitary adenoma</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left optic nerve</td>
<td>3*</td>
<td>-</td>
<td>3*</td>
<td>2</td>
<td>-</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>Right optic nerve</td>
<td>3*</td>
<td>-</td>
<td>3*</td>
<td>2</td>
<td>-</td>
<td>3*</td>
<td>-</td>
</tr>
</tbody>
</table>

**Imaging datasets utilized for OAR contouring**

Table 2 summarizes the image-sets used for delineation of the OARs. For the cavernous sinus meningioma, although various imaging data-sets were used for delineation of the optic-apparatus, most participants contoured the optic-apparatus on either axial contrast-enhanced MRI T1-weighted (T1+C) or the fused image of MRI T2-weighted with CT (FA2). The chiasm was outlined mostly on the fused image of contrast-enhanced MRI T1-weighted with CT (FA) by two of the participants.

Several imaging datasets were available for the pituitary adenoma case but there was no distinct trend of preferred images for contouring. The fused image of MRI T2-weighted with CT was not available for this case while it was one of the preferred images for the contouring of the optic-apparatus in the cavernous sinus meningioma case.

The data-sets most frequently used for contouring the cochlea for the vestibular schwannoma case were MRI T2-weighted (T2) together with MRI T2-weighted fused with CT (FA2) and CT images. Six participants (50%) outlined this OAR as part of clinical routine practice for the vestibular schwannoma case.

**Variability in OAR Delineation**

Table 1 displays the range of contoured volumes, the volume of AV_{50}, AV_{100} and AV_{100/n} together with the AVI and dosimetrical results. The agreement volumes are shown in Figures S2-S4, in the supplementary material, as transparent three-dimensional surfaces. Figure 5 displays the overlapping contours from the TPS in one respective slice for each structure. Figure 1 illustrate one slice of the agreement maps together with one slice of the cochlea contours in their images where they were contoured.

The first row of Table 1 shows the volumes contoured for each structure. The highest consistency was seen for the right optic nerve while the highest inconsistency was seen for the chiasm, both in case 1. The highest discrepancy in the shape, size and position was observed for the left and right optic tract, case 1, illustrated by a value of 0 for the AVI. The highest consistency was observed for the right optic nerve, case 1, where the AVI was 0.13.

**Variability in Dose Constraints**

The chosen Accepted Tolerance Dose (ATD) to the optic-apparatus was 8 Gy in the majority of plans with 2 plans reporting 9 Gy as the ATD. Seven of the 12 participating centers reported no information regarding the tolerance doses for the cavernous sinus and pituitary adenoma cases. For the cochlea only 2 centers reported an ATD of 4 Gy.
Figure 1. Illustration of the regions of agreement for (A) left optic nerve, (B) left optic tract, (C) right optic nerve, (D) right optic tract, (E) chiasm for the cavernous sinus meningioma case, (F) left optic nerve, (G) right optic nerve for the pituitary adenoma case and (H) cochlea for the vestibular schwannoma case. White corresponds to complete or highest level of agreement and levels of grey illustrate lower levels of agreement. Black also corresponds to total agreement. Panels H(i-vi) show the six individual contours and the corresponding image used for contouring (H(i-ii) CT image, H(iii) MRI T2 weighted fused with CT, H(iv-vi) MRI T2 weighted images).
Figure 5. Overlapping contours illustrating the variability in delineation of left (A) and right (B) optic nerve, chiasm (C), left and right optic tract (D) (cavernous sinus meningioma), left (E) and right (F) optic nerve (pituitary adenoma) and cochlea (G) (vestibular schwannoma).

Impact of variable contours on doses to OARs

Dosimetric results are showed in Table 1. The sixth row shows the highest dose to the individual structures. For case 1, the optic-apparatus would receive a maximum dose in the range 3-9 Gy. Similar analysis for case 2, shows a range of 2-5 Gy received by the optic-apparatus. The cochlea in case 3 would receive 3-12 Gy. Row 8 of Table 1 displays the highest dose to the AV100/N from each of the dose-plans. For case 1, this analysis results in a maximum dose to at least one structure of the optic-apparatus ranging between 4-13 Gy across all submitted plans. For case 2, the maximum dose to the optic nerves ranged from 5 to 6 Gy. For case 3, the cochlea received a maximum dose of 12 Gy from at least one of the submitted plans.

DISCUSSION

There is growing evidence that the quality of radiation delivery significantly impacts treatment outcomes. As radiotherapy planning transitions towards an inverse-planning approach, the impact of accurate delineation of the target volume as well as OARs inherently increases. Although the practice of delineating target and OAR volumes is commonly accepted for fractionated radiotherapy, these practices are currently quite variable for radiosurgery. In the present study two structures, the left and right optic tract for case 1, display such great variability that the resulting AVI was 0. Even for the structure with greatest agreement, the right optic nerve in the cavernous sinus case, the AVI was still only 0.13.

Several factors appear to contribute to inconsistency in OAR delineation including the use of variable imaging data-sets, variability in the extent of the structure included in the OAR contour and interobserver variability in the OAR contours, even at common central regions of an OAR. The inconsistency in the range of volumes for the cavernous sinus case was greatest for the optic chiasm with volumes ranging between 0.09-0.61 cm³. For the pituitary adenoma case, there was a wide range of contoured volumes for both the left optic nerve (0.06-0.57 cm³) and right optic nerve (0.10-0.54 cm³). For the vestibular schwannoma case, even when the absolute range of volumes contoured for a small organ like the cochlea was relatively small (0.02-0.05 cm³), the AVI was still low.

Participants in this study reported variable ATDs for OARs. This is reflective of the range of tolerance doses reported for OARs in current radiosurgery literature [12,13,14]. Leber et al [12] reports a tolerance dose of 10 Gy for the optic apparatus when single-fraction radiosurgery is delivered while Tishler et al. [13] concluded that the dose to
the optic-apparatus should be limited to a dose of 8 Gy. Based on the variability in OAR contours observed for the optic structures in our study, one could hypothesize that the variable tolerance-doses reported in these prior studies may have resulted from variability in delineation and dose estimation of these structures across studies. Stafford et al. [14] evaluated clinical outcomes of 218 Gamma Knife treatments of benign tumors in the proximity to the optic-apparatus. As most of the patients (73%) received a dose >8 Gy to a part of the optic-apparatus while complications were observed in <2% of the patients, this study suggested the possibility of allowing greater doses to the optic apparatus if it may improve treatment outcome. Again, the results of this study may have been impacted by variability in the OAR contours themselves and emphasizes the need for standardization to gather stronger data about dose tolerances of OARs for radiosurgery. This study also suggests that accurate delineation and dose estimation to OARs may not only enable protection of OARs but also enable adequate treatment of the target.

In our study, the contoured cochlea in the vestibular schwannoma case was subjected to 3-12 Gy. Prior studies have shown that the likelihood of preserving hearing is enhanced if the doses are lower than 4-5.3 Gy [15,16,17,18]. However, recognizing the potential variability in OAR contours that can impact the collection of consistent dosimetric data in conjunction with clinical follow-up data, there is great need for more prospective data-collection with consistent OAR delineation and dosimetric and clinical reporting to help determine the tolerances dose for the OARs.

One could argue that the main shortcoming of this study is the relaxed set of instructions regarding the contouring and planning provided to the participants. However, this study was intentionally designed to collect data that would mirror current clinical practice at the participating centers. The great variability in the selection, naming and how the structures were contoured in this pilot study was predictable to some extent due to the lack of formal guidelines but also emphasizes the current state of OAR contouring in radiosurgery and the great need for standardized guidelines. For a more formalized quantitative contouring study, guidelines for participants will include instructions regarding the imaging datasets to be used, the specific OAR structures relevant for cases and the extent of each structure to be delineated.

CONCLUSION

The results of this study reveal the wide variability in clinical practice of defining, delineating, and utilizing OARs for radiosurgery planning. The inconsistencies in using OAR contours as part of clinical radiosurgery treatment planning, nomenclature used for OARs, and delineation of OAR volumes significantly hinder the ability to compare results across studies or collect data across institutions for the purpose of determining dose tolerances to OARs for radiosurgery treatment. These findings motivate the radiosurgery community to pursue ongoing efforts to establish a standardized approach and develop guidelines for OAR contouring.

ACKNOWLEDGEMENTS

Financial support from the Cancer Research Funds of Radiumhemmet is gratefully acknowledged as well as the cooperation of all participants involved in this study; Department of Stereotactic and Radiation Neurosurgery and Department of Medical Physics at Na Homolce Hospital in Prague Czech Republic, Ian Paddick MSc and Tim Cox MD at Cromwell Gamma Knife Centre in London England, Katherine Hunt BEng MSc MIPERM, Ruth Batty MD PhD FRCR, Dan Connolly MD PhD FRCR, John Yianni MD FRCSI(SN) and Jeremy Rowe MD MA DM FRCSI(SN) of the National Centre for Stereotactic Radiosurgery at Royal Hallamshire Hospital in Sheffield England, Professor Lijun Ma and Professor Penny Sneed at the UCSF Gamma Knife Center in San Francisco USA, Taylor McAdam Vell Neuroscience Institute at Washington Hospital in Fremont California USA, Francisco Li MSc RSO at the Swedish Radiosurgery Center in Seattle USA, Hyun-Ta Chung PhD at the Department of Neurosurgery Seoul National University College of Medicine in Seoul Korea, Monique van Prooijen PhD at the Princess Margaret Cancer Centre in Toronto Canada, Chryssa Paraskevopoulou MSc of Hygieia Hospital in Athens Greece, Bente Sandevi Skeie MD PhD and Jan Heggdal MSc at the Department of Oncology and Medical Physics at Haukeland University Hospital in Bergen Norway, David Schlesinger PhD, David Larson MD and Kevin Orcutt MD from the Department of Radiation Oncology at the University of Virginia in Charlottesville Virginia USA. Technical assistance of Björn Somell at Elekta Instrument AB in Stockholm is also gratefully acknowledged.

INFORMED CONSENT

All patients included in this study have given a written informed consent.
REFERENCES


Figure S2. 3D illustrations for the OAR delineated in the cavernous sinus meningioma case. Left column shows the AV100 (blue) and AV50/N (green) and the right column shows the AV50 (blue) together with the encompassing volume (green) for (A) and (B) left optic nerve, (C) and (D) left optic tract, (E) and (F) right optic nerve, (G) and (H) right optic tract, (I) and (J) chiasm.
Figure S3. 3D illustrations for the OAR delineated in the pituitary adenoma case. Left column shows the AV$_{100}$ (blue) and AV$_{100/N}$ (green) and the right column shows the AV$_{50}$ (blue) together with the encompassing volume (green) for (A) and (B) left optic nerve, (C) and (D) right optic nerve.

Figure S4. 3D illustrations for the OAR delineated in the vestibular schwannoma case. Left figure shows the AV$_{100}$ (blue) and AV$_{100/N}$ (green) and the right figure shows the AV$_{50}$ (blue) together with the encompassing volume (green) for (A) and (B) cochlea.
Research Article

Radiobiological Framework for the Evaluation of Stereotactic Radiosurgery Plans for Invasive Brain Tumours

Helena Sandström,1 Alexandru Dasu,2 and Iuliana Toma-Dasu1

1 Medical Radiation Physics, Stockholm University and Karolinska Institutet, 171 76 Stockholm, Sweden
2 Department of Radiation Physics UHL, County Council of Östergötland, Linköping University, 581 85 Linköping, Sweden

Correspondence should be addressed to Iuliana Toma-Dasu; iuliana.livia.dasu@ki.se

Received 21 August 2013; Accepted 18 October 2013

Academic Editors: S. T. Chao, A. Goussia, G. Metro, and S. Mohanam

Copyright © 2013 Helena Sandström et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study presents a radiobiological formalism for the evaluation of the treatment plans with respect to the probability of controlling tumours treated with stereotactic radiosurgery accounting for possible infiltrations of malignant cells beyond the margins of the delineated target. Treatments plans devised for three anaplastic astrocytoma cases were assumed for this study representing cases with different difficulties for target coverage. Several scenarios were considered regarding the infiltration patterns. Tumour response was described in terms of tumour control probability (TCP) assuming a Poisson model taking into account the initial number of clonogenic cells and the cell survival. The results showed the strong impact of the pattern of infiltration of tumour clonogens outside the delineated target on the outcome of the treatment. The treatment plan has to take into account the existence of the possible microscopic disease around the visible lesion; otherwise the high gradients around the target effectively prevent the sterilisation of the microscopic spread leading to low probability of control, in spite of the high dose delivered to the target. From this perspective, the proposed framework offers a further criterion for the evaluation of stereotactic radiosurgery plans taking into account the possible infiltration of tumour cells around the visible target.

1. Introduction

The aim of radiation therapy is to stop the tumour growth process with sparing of the normal tissues nearby. For stereotactic radiosurgery (SRS) this is achieved by delivering a highly conformal dose distribution to the target in one fraction. The relatively steep dose falloff around the target ensures the sparing of the normal tissue and/or the critical structures near the target and this is the core of the SRS concept. The evaluation of plans is currently performed as a function of the conformity of therapeutic isodoses to the defined target and the gradients outside the target. This approach intrinsically assumes that tumour cells are confined to the target volume and that there are no infiltrations in the normal tissues around this target or that the impact of the possible infiltrations outside the delineated target on the probability of eradication of the tumour is negligible. However, several of the brain tumours commonly treated with SRS are invasive and therefore the existence of tumour cells outside the tumour lesions that could be identified in diagnostic images cannot be excluded [1, 2].

From this perspective, the evaluation of the plans should be performed not only from purely geometrical and physical points of view, but also from a radiobiological perspective taking into account the invasiveness of the tumours that have to be treated and the distribution of tumour cells in and around the target. Therefore it is the aim of this paper to introduce a radiobiological formalism for the evaluation of the treatment plans with respect to the probability of controlling tumours treated with SRS.

2. Materials and Methods

2.1. Patient Material and Target Definition. Three representative cases of recurrent anaplastic astrocytoma have been selected from a series of cases treated with Leksell Gamma Knife Perfexion (Figure 1). The treatment plans were calculated for a prescribed dose to the target of 16 Gy at the
50% isodose. The dose distributions were exported from the treatment planning system and used for calculations together with the structures. Dose matrices were exported from the treatment plans with the same transversal resolution as the structure matrices. The interslice resolution of the structure matrices is given by the imaging method used in each case. All dose matrices were redefined to have the same number of slices and interslice resolution as the structure matrices.

The three panels in Figure 1 illustrate the cases chosen for this study. Case 1 (Figure 1(a)) shows a plan for the anaplastic astrocytoma with poor conformity. The conformity was quantified and expressed as conformity index, defined as the ratio of the volume of the target covered by the prescribed isodose volume and the total target volume [3]. The conformity index (CI) for case 1 was 0.77. A plan for which a much higher CI has been achieved is presented in Figure 1(b), hereby described as case 2. The CI for case 2 was 0.96. Figure 1(c) illustrates an intermediate situation, case 3, for which the plan leads to a CI of 0.91.

Each of the cases in Figure 1 is not only showing a different coverage of the target but also a different irradiation of the tissue nearby the target.

In order to account for the invasiveness of the anaplastic astrocytoma clonogenic cells outside the planned target volume, several scenarios regarding infiltration of tumour cells have been considered. Thus, in one scenario it has been assumed that there is no infiltration around the delineated tumour volume and that the cell density is the same in all the voxels of the target. The other two scenarios assumed that tumour clonogenic cells exist outside the delineated area, either as a continuously decreasing function of the distance to the target or assuming a more heterogeneous pattern of infiltration outside the target (Figure 2). Although there are several modelling studies of the tumour margin diffusion and infiltration [4], the information regarding the infiltration pattern of the astrocytoma cells is rather scarce. In absence of histologically validated models for the infiltrations, a simple continuous function describing the decrease of the density of clonogenic cells outside the target with distance as \( f(d) = \exp(-d) \) was assumed. For modelling the stochastic character of invasiveness suggested by some studies [4], a second scenario was also assumed in which the continuous function describing the decrease of the density of clonogenic cells outside the target with distance was coupled with a random distribution of the clonogens in the voxels outside the target.

In both scenarios, as astrocytomas are highly infiltrative types of gliomas, similar to glioblastomas, the maximal distance for infiltration was based on a study by Yamahara et al. [5] comparing examined autopsy brains and MR images for glioblastoma in which a peripheral tumour boundary infiltration of 6–14 mm was found. Thus, the maximum distance at which the astrocytoma cells could be found outside the borders of the target was considered to be 10 mm, the average value in the study by Yamahara et al. [5].

2.2. Radiobiological Model. In order to assess the influence of the tumour cells invasiveness outside the target on the treatment outcome, the response of the tumour to radiosurgery was described in terms of the tumour control probability (TCP) assuming a Poisson model taking into account the initial number of clonogenic cells and the cell survival. Given a dose distribution with doses \( D_i \) to voxels \( i \) in the patient, the tumour control probability is described by

\[
TCP = \exp \left( -\sum_{i=1}^{n} \rho_0(V_i) V_i e^{-\alpha D_i - \beta D_i^2} \right),
\]

where \( n \) is the total number of voxels, \( \rho_0(V_i) \) is the initial density of clonogenic cells in voxel \( i \), \( V_i \) is the volume of voxel \( i \), and \( D_i \) is the dose delivered to the cells in voxel \( i \). Equation (1) assumes the LQ model for cell survival [6] with parameters \( \alpha \) and \( \beta \). It has to be mentioned that the general expression in (1) could be used for response calculation irrespective of the size or shape of the tumour and any given distribution of doses \( D_i \).

The radiobiological parameters used in the calculations were \( \alpha = 0.24 \text{ Gy}^{-1}, \beta = 0.03 \text{ Gy}^{-2} \), and \( \alpha / \beta = 8.31 \text{ Gy} \) as reported in the review by Malaise et al. [7] for glioblastoma. It was further assumed that the cell population in the target
Figure 2: Delineated targets from Figure 1 and patterns of infiltration assumed in this study. Cell density outside the target is assumed to decrease as function of the distance to the target (upper panels) or have a more heterogeneous pattern of infiltration (lower panels). The colour scale indicates the relative density of clonogens, from 1 inside the target volume to zero outside the target if the distance from the border of the target exceeded 10 mm.

is of the order of $10^6$ cells in order to have a normalised slope of the TCP curve similar to that of gliomas.

### 3. Results and Discussion

The distribution of cell survival in and around the target for each infiltration scenario considered in this study is shown in Figure 3.

The variation of cell survival within the target indicates that if the malignant cells are confined to the target, the high doses delivered to it are enough to sterilise all cells. However, if tumour cells may be found outside the delineated target, the steep dose falloff outside the target leads to lower cell killing in the regions where malignant cells may be infiltrating. This translates into very high TCP values for tumour cells confined to the treated volume and very poor outcome for infiltrations that are not accounted for during planning.

One might expect that the reduced cell kill might balance the decreased density of the infiltrating cells, but in reality the volume around the target in which infiltrating clonogenic cells might be encountered represents a volume equal to if not larger than the target (e.g., 10 mm infiltration distance around a 2 cm diameter spherical target represents a volume 7 times larger than the target itself). Hence, the relatively large volume represented by the margin effectively leads to rather high survival that is reflected in very low TCP. Thus, the calculated tumour control probability for the three cases shown in Figure 1 is dropping from 100% when no infiltration is assumed to 0% for the areas of infiltration illustrated in Figure 2 (Table 1).

These results support the idea that radiobiological outcome from SRS treatments could depend on the potential for infiltration of the cells in the treated tumour and the way this is taken into account at the stage of target delineation. Thus, the existence of the possible microscopic disease around the lesion visible on diagnostic images might effectively ruin the probability of controlling the tumour with SRS. This is important for many brain tumours that are candidates for SRS. Indeed, a study by Yamahara et al. [5] compared magnetic resonance (MR) images with pathological samples of glioblastoma multiforme (GBM) and showed that tumour cells may exist as far as 14 mm outside the tumour boundary.

<table>
<thead>
<tr>
<th>No infiltration of clonogens outside the target</th>
<th>Continuous decrease of density of clonogens outside the target</th>
<th>Heterogeneous decrease of density of clonogens outside the target</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>Case 2</td>
<td>Case 3</td>
</tr>
<tr>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 1: Tumour control probability for the three plans in Figure 1 and the corresponding scenario for the invasiveness of the astrocytoma clonogenic cells.


determined on MR images. Similarly, Pirzkall et al. [8] verified the validity of magnetic resonance spectroscopy (MRSI) in defining the extent of glioma infiltration and suggested an addition of 2-3 cm margin to the gross tumour volume. Another study by the same group showed that the infiltration is heterogeneous and therefore a nonuniform margin might be needed [9].

The results in this study have been obtained under the assumption that cell density is constant within the delineated target. However, the delineated target shows heterogeneous uptake of contrast agents used to identify lesions in diagnostic images, which may also indicate heterogeneous cell density. Investigating the relationship between contrast uptake and cell density is however beyond the purpose of the present study that aimed to construct a radiobiological formalism for SRS evaluation taking into account possible infiltrations of malignant cells outside the target. Nevertheless, our results indicate that the existence of undetected tumour cells outside the volume receiving therapeutic doses might effectively lead to treatment failure as they could regrow the tumour following the treatment. It could be that, in some cases, when for example, some margins are included in the target, the doses outside it might be enough to sterilise the lower number of tumour cells infiltrated in the surrounding normal tissue. In others, however, the high gradients around the target might effectively prevent the sterilisation of the microscopic disease leading to a recurrence near the treated volume. From this perspective, further studies on the distribution of malignant cells in and around the lesions visible on diagnostic images are therefore warranted.

The simulations in this study have been performed with the LQ model. While the validity of the LQ model has been debated in recent years for SRS treatments employing large fractional doses delivered to the target [10, 11], it is important to recognise that the focus of this study has been on the effects in regions around the target that receive lower than therapeutically prescribed doses. For these doses, the validity of the LQ model is not a matter of debate and therefore the significance of the results cannot either be debated. They clearly indicate that if cell kill outside the target does not counterbalance the infiltrating cells, the outcome of the treatment might be poorer than expected.

The results of the present study therefore highlight the importance of target delineation and warrant further studies regarding patterns of failure for tumours treated with SRS and the relationship to the potential for infiltration of the treated tumour.

4. Conclusions

A radiobiological framework for the evaluation of treatment plans for invasive brain tumours was developed, taking into
account the invasiveness of the tumour into the surrounding normal tissues. This offers a further criterion for the evaluation of stereotactic radiosurgery plans besides the conformity of therapeutic isodoses to the defined target and the gradients outside the target.

**Conflict of Interests**

All the authors declare they have no conflict of interests.

**Acknowledgments**

Financial support from the Cancer Research Funds of Radiumhemmet is gratefully acknowledged. The authors would also like to thank Håkan Nordström and Jonas Johansson from Elekta Instrument AB (Stockholm, Sweden) for stimulating discussions and support, Hidefumi Jokura from Jiro Suzuki Memorial Gamma House, Furukawa Seiryo Hospital (Osaki, Japan), and Department of Neurosurgery, Tohoku University School of Medicine (Sendi, Japan), for providing the patient data, and Pierre Barsoum from Karolinska University Hospital (Stockholm, Sweden) for technical support.

**References**


V
To fractionate or not to fractionate? That is the question for the radiosurgery of hypoxic tumors

Laboratory investigation

IULIANA TOMA-DASU, Ph.D., HELENA SANDSTRÖM, M.Sc., PIERRE BARSOUM, M.Sc., and ALEXANDRU DASU, Ph.D.

1Medical Radiation Physics, Stockholm University and Karolinska Institute; 2Department of Medical Physics, Karolinska University Hospital, Stockholm; and 3Departments of Radiation Physics and Medical and Health Sciences, Linköping University, Linköping, Sweden

Object. This study aimed to investigate the impact of tumor hypoxia on treatment outcome for metastases commonly treated with radiosurgery using 1 fraction of radiation and the potential gain from reoxygenation if the treatment is delivered in a few radiation fractions.

Methods. In silico metastasis-like radiosurgery targets were modeled with respect to size, density of clonogenic cells, and oxygenation. Treatment plans were produced for the targets using Leksell GammaPlan, delivering clinically relevant doses and evaluating the tumor control probability (TCP) that could be expected in each case. Fractionated schedules with 3, 4, and 5 fractions resulting in similar biological effective doses were also considered for the larger target, and TCP was determined under the assumption that local reoxygenation takes place between fractions.

Results. The results showed that well-oxygenated small- and medium-size metastases are well controlled by radiosurgery treatments delivering 20 or 22 Gy at the periphery, with TCPs ranging from 90% to 100%. If they are moderately hypoxic, the TCP could decrease to 60%. For large metastases, the TCPs from single-fraction treatments ranged from 0% to 19%, depending on tumor oxygenation. However, for fractionated treatments, the TCP for hypoxic tumors could significantly increase up to 51%, if reoxygenation occurs between fractions.

Conclusions. This study shows that hypoxia worsens the response to single-fraction radiosurgery, especially for large tumors. However, fractionated therapy for large hypoxic tumors might considerably improve the TCP and might constitute a simple way to improve the outcome of radiosurgery for patients with hypoxic tumors.

D Ellevering a highly conformal dose distribution to the target while sparing the normal tissues surrounding it is one of the central principles of stereotactic radiosurgery. Nevertheless, the distribution of the dose to the tissue is determined by the physics of radiation interaction with the matter and the technological characteristics of the radiation delivery system. Consequently, the dose to the target is often dictated by the tolerance of the normal tissue. Furthermore, the inverse correlation between the volume of the target and the dose that can be delivered to it usually limits the target sizes that can be irradiated with stereotactic techniques. This correlation also leads to the paradox that larger tumors, presumably containing more cells, are treated with lower doses than are smaller tumors with fewer cells.

The success of stereotactic radiosurgery also depends on clinical and radiobiological factors. Delivering the dose in 1 fraction addresses 2 of the classic Rs of radiobiology, repair and repopulation: delivering a single fraction will not allow time for the tumor to repair sublethal damage or to proliferate and thus to increase the target population through repopulation. Given the sizes of the doses usually delivered, redistribution will not likely influence the outcome. However, oxygenation status of the tumors to be treated will likely influence the outcome, and treatments employing 1 or very few fractions are less influenced by reoxygenation. Indeed, delivering the dose in 1 or a few fractions, as in stereotactic body radiation therapy, prevents slow reoxygenation caused by preferentially killing the well-oxygenated cells at the periphery of the tumor and thus decreasing oxygen consumption and making oxygen available to the hypoxic cells located in the core of the tumor, as described by Thomlinson and Gray. However, short-term local reoxygenation due to opening of temporarily closed blood vessels, as initially postulated by Brown and later demonstrated by Chaplin et al., a process independent of radiation delivery, could change cellular radiosensitivity during multifraction delivery. This study aimed to explore the impact of tumor hypoxia on the outcome of radiosurgery to treat metastases and the potential of fast

Abbreviations used in this paper: BED = biological effective dose; HF = hypoxic fraction; HI = heterogeneity index; TCP = tumor control probability.
To fractionate or not to fractionate?

reoxygenation to improve the results of fractionated dose delivery for this type of treatment.

**Methods**

The impact of hypoxia and reoxygenation was studied with in silico tumor models with heterogeneous oxygenations. Three sizes of radiosurgery targets were assumed: 1, 2, and 3 cm in diameter. Also, 3 possible oxygenation states were assumed for the tumors (Fig. 1). They could be well oxygenated, or in the case of larger tumors, they could have a moderately hypoxic or a more hypoxic core. The oxygen distributions calculated for each region accounted for relevant distributions of intervessel distances and the physical processes of oxygen diffusion and consumption, as extensively described in previous publications. The density of clonogenic cells was assumed to be a uniform 10⁶ cells/cm³ throughout the tumors.

For each target, a plan was made in the Leksell GammaPlan version 10.1.1 (Elekta AB) with the 50% isodose covering the target (Fig. 2). The dose prescription in each plan was target size dependent, with 22 Gy to the 50% isodose for the 1-cm target and 20 Gy and 18 Gy to the 2-cm and 3-cm targets, respectively (Fig. 3). Table 1 lists other dosimetric parameters of interest for each plan, including the gradient index, the Paddick conformity index, and the heterogeneity index HI:

\[
HI = \frac{D_{\text{max}} - D_{\text{min}}}{D_{\text{mean}}},
\]

where \(D_{\text{max}}\) is the maximum dose, \(D_{\text{min}}\) is the minimum dose, and \(D_{\text{mean}}\) is the mean dose to the target volume.

**Fig. 1.** Sections through the central plane of tumor models used in the study, showing examples of the different oxygenation states considered. **A:** A well-oxygenated small tumor, 1 cm in diameter (0.52 cm³ volume); the HF, calculated as the percentage of values with \(pO_2 < 5 \text{ mm Hg}\), is 0%. **B:** A moderately hypoxic (HF = 4%) medium-size tumor, 2 cm in diameter (4.19 cm³ volume). **C:** A hypoxic (HF = 10%) large-size tumor, 3 cm in diameter (14.14 cm³ volume).

Tumor response was assessed as tumor control probability (TCP), calculated as

\[
TCP = \exp \left\{ - \sum_{i=1}^{N_{\text{vox}}} N_i \left[ SF\left(d_i, p_i\right) \right] \right\},
\]

where \(N_{\text{vox}}\) is the number of calculation voxels in the tumor, \(N_i\) is the number of cells in voxel \(i\), and \(SF(d_i, p_i)\) is the cell survival in voxel \(i\) dependent on dose \(d_i\) and oxygen tension \(p_i\). Cell survival was calculated with the linear quadratic model, using the formalism described in Toma-Dasu and Dasu. Thus, cell survival \(SF_i\) in each voxel receiving dose \(d_i\) is described by

\[
SF(d_i, p_i) = \exp \left[ -\frac{\alpha}{\text{OMF}(p_i)} d_i - \frac{\beta}{\text{OMF}^2(p_i)} d_i^2 \right],
\]

where \(\alpha\) and \(\beta\) are the linear quadratic parameters relevant for oxic cells and OMF(\(p_i\)) are oxygen tension–dependent modification factors:

\[
\text{OMF}(p_i) = \frac{k + OER_{\text{max}} p_i}{k + OER_{\text{max}} p_i},
\]

where OER_{\text{max}} is the maximum protection achieved in the absence of oxygen and \(k\) is a reaction constant as described by Alper and Howard-Flanders. The following oxic parameters were assumed for calculation: \(\alpha = 0.35\) Gy⁻¹ and \(\alpha/\beta = 10\) Gy.
Fractionated schedules with 3, 4, and 5 fractions leading to the same biological effective doses (BEDs) to the tumor (i.e., for α/β = 10 Gy) were also considered for the larger target (Table 2). In this case, TCP was calculated assuming that fast local reoxygenation takes place between fractions.

**Results**

The simulations showed that well-oxygenated small- and medium-size metastases could be very well controlled by radiosurgery treatments delivering a single fraction of 20 Gy or 22 Gy at the periphery, with TCPs ranging from 90% to 100%. However, for moderately hypoxic tumors, the TCP would decrease to 60%. In contrast, large metastases were difficult to control with single-fraction regimes, with TCPs ranging from 0% to 19%, depending on the assumed oxygenation state. Thus, it appears that hypoxia worsens the response to single-fraction radiosurgery, especially for large tumors.

With fractionated treatments for hypoxic tumors, interfraction reoxygenation resulted in significantly increased TCPs, up to 51% (Table 3). Better results are obtained for oxic and even moderately hypoxic tumors, but even large, severely hypoxic tumors could be controlled with fractionated regimes. Furthermore, for the latter the TCP increases with the number of fractions, highlighting the importance of the number of opportunities for reoxygenation to improve the possible outcome. These results therefore emphasize the importance of using pretreatment tumor oxygenation status to choose the best treatment option for radiosurgery patients.

**Discussion**

Tumor hypoxia is an important determinant of outcome in radiation therapy, as the absence of oxygen could render the cells resistant to radiation and could also contribute to the selection of more aggressive tumor cell phenotypes. Several proposals have been made to

---

**Fig. 2.** Dose distributions in the central plane for each target size illustrated in Fig. 1, with the 50% isodose covering the target.

**Fig. 3.** Dose distributions in the central plane for each target size illustrated in Fig. 1, with 22 Gy to the 50% isodose for the 1-cm target (A), 20 Gy for the 2-cm target (B), and 18 Gy for the 3-cm target (C).
To fractionate or not to fractionate?

TABLE 1: Treatment plan dosimetric parameters for the targets considered in this study

<table>
<thead>
<tr>
<th>Parameter</th>
<th>10-mm Target</th>
<th>20-mm Target</th>
<th>30-mm Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-mm shots</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8-mm shots</td>
<td>1</td>
<td>3</td>
<td>14</td>
</tr>
<tr>
<td>16-mm shots</td>
<td>0</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>composite shots</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>grid size (mm)</td>
<td>0.7</td>
<td>1.2</td>
<td>1.6</td>
</tr>
<tr>
<td>coverage</td>
<td>99%</td>
<td>99%</td>
<td>98%</td>
</tr>
<tr>
<td>selectivity</td>
<td>92%</td>
<td>97%</td>
<td>97%</td>
</tr>
<tr>
<td>gradient index</td>
<td>3.20</td>
<td>2.56</td>
<td>2.70</td>
</tr>
<tr>
<td>Paddick conformity index</td>
<td>0.91</td>
<td>0.96</td>
<td>0.95</td>
</tr>
<tr>
<td>HI</td>
<td>0.72</td>
<td>0.82</td>
<td>0.87</td>
</tr>
</tbody>
</table>

counteract the hypoxic cells in conventional radiotherapy, although they have been slow to enter clinical practice.29 The most recent proposals for antihypoxic strategies aiming to increase the doses to the hypoxic foci in tumors are quite promising, and several studies have been initiated to investigate their clinical potential.37,39,40

Hypoxia is not usually taken into account for stereotactic radiosurgery, in part because of the large doses that could be delivered to the target while keeping irradiation of normal tissues within tolerances. Nevertheless, the results of this study show that hypoxia could negatively affect the results from single-fraction stereotactic radiosurgery, especially for larger tumors, for which normal tissue imposes limitations on the maximum dose that can be delivered to the target. Given that the options available for patients with large inoperable cranial tumors are usually limited to palliative treatments or treatments with conventional fractionations, it is worthwhile to explore whether they could benefit from stereotactic radiation treatments. The present study shows that allowing local reoxygenation of the tumors by fractionating the treatment could significantly improve the outcome for patients with large tumors by taking advantage of the favorable dose distributions that could be delivered with radiosurgery techniques.

These results were obtained under the constraint that the fractionated treatments deliver the same BEDs to the edge of the target. The corresponding effect to the normal brain decreases with increasing fractionation, as shown in the last column in Table 2, due to the lower values of α/β

TABLE 2: Fractionated schedules with 3, 4, and 5 fractions leading to the same BEDs to the tumor, BED10, as 18 Gy delivered in a single fraction, and the corresponding BED to normal tissue, BED2

<table>
<thead>
<tr>
<th>Total Dose D (Gy)</th>
<th>Dose per Fraction d (Gy)</th>
<th>No. of Fractions n</th>
<th>BED10 (Gy)</th>
<th>BED2 (Gy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>18.0</td>
<td>18.0</td>
<td>1</td>
<td>50.4</td>
<td>180.0</td>
</tr>
<tr>
<td>26.7</td>
<td>8.9</td>
<td>3</td>
<td>50.4</td>
<td>145.3</td>
</tr>
<tr>
<td>29.2</td>
<td>7.3</td>
<td>4</td>
<td>50.4</td>
<td>135.4</td>
</tr>
<tr>
<td>31.1</td>
<td>6.2</td>
<td>5</td>
<td>50.4</td>
<td>127.7</td>
</tr>
</tbody>
</table>

Our results were obtained from theoretical simulations and may therefore be influenced by the inherent limitations of the method. The simulations focused on the direct effect of radiation on the tumor cells, whereas other mechanisms may also be involved in cell death. For example, the immune system may be responsible for removing a few logs of cells.14 Although accurate numbers are not available for this process, it was partly included in our simulation by assuming a low density of clonogens per unit volume. Furthermore, tumor cells depend on the supply of nutrients through the vascular network. Tumor vasculature, although not directly associated with radiosurgery, has long been considered a possible target to increase cell death.12,13 A high-dose fraction in the range of 8–11 Gy may induce endothelial apoptosis of the tumor vasculature, thus indirectly increasing cell death by abolishing nutrient supply,3 much like vascular targeting. This effect is in addition to direct cell death by radiation, and it could increase the TCPs above those given in Table 3, as suggested by Park et al.32 The additional effect of vasculature damage could decrease with increased fractionation, because the fractional dose is lower than the threshold for inducing vascular effects. Nevertheless, the effect of vascular damage changes the baseline TCP values for extremely hypofractionated schedules, and hypoxia could be regarded as an additional factor able to modulate response to radiation therapy. In this respect, the results of this study also agree with those of Carlson et al.,6 who showed that extreme hypofractionation should be avoided for hypoxic tumors because reoxygenation possibilities are severely reduced during the treatment course.

Another possibility that was not included in this simulation was that cellular density might vary with tumor.

TABLE 3: Tumor control probabilities for a large tumor under 2 oxygenation conditions*

<table>
<thead>
<tr>
<th>Total Dose D (Gy)</th>
<th>No. of Fractions n</th>
<th>Moderately Hypoxic (HF = 4%)</th>
<th>Hypoxic (HF = 10%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>18.0</td>
<td>1</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>26.7</td>
<td>3</td>
<td>46%</td>
<td>11%</td>
</tr>
<tr>
<td>29.2</td>
<td>4</td>
<td>51%</td>
<td>35%</td>
</tr>
<tr>
<td>31.1</td>
<td>5</td>
<td>51%</td>
<td>43%</td>
</tr>
</tbody>
</table>

For the treatment schedules using ≥ 1 fraction, TCP was calculated assuming that local reoxygenation occurs between fractions.
oxygenation. Severely hypoxic tumors could have fewer viable cells toward their center, as necrosis would set in. In this case, an increased dose to the center would be delivered to a sparsely populated region of the tumor. From this perspective, the models in this study and TCP results should be regarded as worst-case scenarios regarding the impact of hypoxia and improvements that could be obtained in clinical settings. Nevertheless, this study shows the significant improvements that could be obtained through fractionation even for the most unfavorable intratumor cell distributions (Table 3).

Finally, the validity of the linear quadratic model for the large doses employed in radiosurgery is an ongoing debate. However, the present study focused on the response of mixed cell populations containing a large hypoxic subpopulation for which the linear quadratic model is able to accurately describe the response, so the results should be relevant for clinically used dose prescriptions to cranial targets treated with stereotactic radiosurgery.

An important question is whether predictions from theoretical simulations are compatible with clinical experience. Several analyses show that local control for single-fraction radiosurgery decreases when the dose delivered to the tumor margin decreases or when tumor size increases. Furthermore, the results reported by these analyses are compatible with the theoretical TCP results in the present study; vascular and immunologic effects are probably responsible for any differences. Chang et al. reported local control rates at 1 and 2 years, respectively, of 86% and 78% for lesions < 1 cm, and 56% and 24% for lesions > 1 cm, treated with a minimum peripheral dose of 20 Gy (BED$_{90}$ = 60 Gy$_{90}$). Similarly, Hasegawa et al. reported local control rates of 84% and 77%, respectively, for tumors < 4 cm (2 cm equivalent diameter) at 1 and 2 years after treatment with a mean marginal dose of 18.5 Gy (BED$_{90}$ = 52.7 Gy$_{90}$). Vogelbaum et al. reported 1-year control rates of 85% in tumors < 2 cm treated with 24 Gy (BED$_{90}$ = 81.6 Gy$_{90}$), 49% in tumors of 2–3 cm treated with 18 Gy (BED$_{90}$ = 50.4 Gy$_{90}$), and 45% in tumors > 3 cm treated with 15 Gy (BED$_{90}$ = 37.5 Gy$_{90}$). More recently, Fokas et al. reported 73% local control at 1 year after 20 Gy in 138 patients with tumors with a median diameter of 1.5 cm.

In contrast, fractionated radiosurgery appears to lead to improved local control, especially for larger tumors, in line with results of the present study. Aoyama et al. reported 81% local control at 1 year in 87 patients with tumors with a median equivalent diameter of 1.8 cm treated with a median dose of 35 Gy in 4 fractions (BED$_{90}$ = 65.6 Gy$_{90}$). Narayana et al. reported 70% local control at 1 year after 30 Gy in 5 fractions (BED$_{90}$ = 48 Gy$_{90}$) in 20 patients with 1.9-cm tumors. Higuchi et al. analyzed 43 patients with tumors > 3 cm and reported 76% local control at 1 year after 30 Gy in 3 fractions (BED$_{90}$ = 60 Gy$_{90}$). Fokas et al. reported 1-year local control rates of 75% and 71% in 122 patients treated either with 35 Gy in 7 fractions (BED$_{90}$ = 52.5 Gy$_{90}$) for tumors of about 1.6 cm or with 40 Gy in 10 fractions (BED$_{90}$ = 56 Gy$_{90}$) for tumors of about 2.2 cm, respectively. Rajakesari et al. reported 56% local control at 1 year in 70 patients with a median tumor diameter of 1.7 cm treated mainly with 25 Gy in 5 fractions (BED$_{90}$ = 37.5 Gy$_{90}$). Minniti et al. reported 1- and 2-year local control rates of 88% and 75%, respectively, for 135 patients treated with 36 Gy in 3 fractions (BED$_{90}$ = 79.2 Gy$_{90}$) for tumors < 2 cm and with 27 Gy in 3 fractions (BED$_{90}$ = 51.3 Gy$_{90}$) for tumors > 2 cm. Fractionated treatments also appear to lead to lower toxicity rates than single-fraction treatments, also in line with the proposals of the present study. Although only a randomized trial will show whether fractionated radiosurgery may improve the outcome of larger tumors with less toxicity, this in silico analysis nevertheless appears to support this hypothesis. Further theoretical and practical exploration of the potential of fractionated stereotactic radiosurgery is therefore warranted.

Conclusions

This study showed that hypoxia worsens the response to single-fraction radiosurgery, especially for large tumors. Fractionated therapy for large hypoxic tumors might considerably improve the TCP. These results therefore suggest that determining pretreatment tumor oxygenation status and considering fractionation for the patients with hypoxic tumors might improve the outcome of radiosurgery. This may be a practical and relatively simple way to individualize treatment for radiosurgery patients.

Disclosure

Financial support from the Cancer Research Funds of Radiumhemmet, Linköping University, and the County Council of Östergötland is gratefully acknowledged. The authors report no conflict of interest concerning the materials or methods used in this study or the findings in this paper.

Author contributions to the study and manuscript preparation include the following: Conception and design: Toma-Dasu, Dasu. Acquisition of data: Sandström, Barsoum. Analysis and interpretation of data: all authors, Drafting the article: Toma-Dasu, Dasu. Critically revising the article: all authors. Reviewed submitted version of manuscript: all authors. Approved the final version of the manuscript on behalf of all authors: Toma-Dasu.

References

To fractionate or not to fractionate?


Manuscript submitted June 30, 2014. Accepted August 4, 2014.
Please include this information when citing this paper: DOI: 10.3171/2014.8.GKS141461.
Address correspondence to: Iuliana Toma-Dasu, Ph.D., Medical Radiation Physics, Karolinska Institutet, Box 260, 171 76 Stockholm, Sweden. Email: iuliana.livia.dasu@ki.se.