

RESEARCH

Open Access



CrossMark

# Variability in spine radiosurgery treatment planning – results of an international multi-institutional study

André Toussaint<sup>1\*</sup>, Anne Richter<sup>1</sup>, Frederick Mantel<sup>1</sup>, John C. Flickinger<sup>2</sup>, Inga Siiner Grills<sup>3</sup>, Neelam Tyagi<sup>3</sup>, Arjun Sahgal<sup>4</sup>, Daniel Letourneau<sup>5</sup>, Jason P. Sheehan<sup>6</sup>, David J. Schlesinger<sup>6</sup>, Peter Carlos Gerszten<sup>2</sup> and Matthias Guckenberger<sup>1,7</sup>

## Abstract

**Background:** The aim of this study was to quantify the variability in spinal radiosurgery (SRS) planning practices between five international institutions, all member of the Elekta Spine Radiosurgery Research Consortium.

**Methods:** Four institutions provided one representative patient case each consisting of the medical history, CT and MR imaging. A step-wise planning approach was used where, after each planning step a consensus was generated that formed the basis for the next planning step. This allowed independent analysis of all planning steps of CT-MR image registration, GTV definition, CTV definition, PTV definition and SRS treatment planning. In addition, each institution generated one additional SRS plan for each case based on intra-institutional image registration and contouring, independent of consensus results.

**Results:** Averaged over the four cases, image registration variability ranged between translational 1.1 mm and 2.4 mm and rotational 1.1° and 2.0° in all three directions. GTV delineation variability was 1.5 mm in axial and 1.6 mm in longitudinal direction averaged for the four cases. CTV delineation variability was 0.8 mm in axial and 1.2 mm in longitudinal direction. CTV-to-PTV margins ranged between 0 mm and 2 mm according to institutional protocol. Delineation variability was 1 mm in axial directions for the spinal cord. Average PTV coverage for a single fraction 18 Gy prescription was  $87 \pm 5\%$ ;  $D_{min}$  to the PTV was  $7.5 \pm 1.8$  Gy averaged over all cases and institutions. Average  $D_{max}$  to the PRV\_SC (spinal cord + 1 mm) was  $10.5 \pm 1.6$  Gy and the average Paddick conformity index was  $0.69 \pm 0.06$ .

**Conclusions:** Results of this study reflect the variability in current practice of spine radiosurgery in large and highly experienced academic centers. Despite close methodical agreement in the daily workflow, clinically significant variability in all steps of the treatment planning process was demonstrated. This may translate into differences in patient clinical outcome and highlights the need for consensus and established delineation and planning criteria.

**Keywords:** Spine radiosurgery, Vertebral metastases, Delineation, Planning variability

## Introduction

Cancer patients with systemic tumor spread frequently develop skeletal metastases. Many of these occur within the spine [1–3]. Radiotherapy (RT) is an established and recommended component of the multidisciplinary treatment of spine metastases with regard to prevention of pathologic fractures or neurological deficits and pain palliation [4–6]. Conventional RT with low dose per

fraction is effective but has been shown to achieve only a rather short duration of pain response of only 3–6 months (median) [7]. Improvements in systemic treatment efficacy have prolonged survival in many cancer patients. Also, validated overall survival scores have been described for patients treated with spine metastases that has enabled identification of patient subgroups with longer survival [8–10].

Spine Stereotactic Radiosurgery (SRS), sometimes also referred to as Stereotactic Body Radiation Therapy (SBRT) to the spine, has been demonstrated to result in promising long term local control and pain palliation

\* Correspondence: toussaint\_a@ukw.de

<sup>1</sup>Department of Radiation Oncology, University of Wuerzburg, Wuerzburg, Germany

Full list of author information is available at the end of the article

with low toxicity rates [11–16]. The use of SRS as an alternative treatment to conventional RT has been increasing rapidly. However, there still remains significant variability and little consensus exist regarding target volume contouring and treatment planning [17–19].

The aim of this study was to quantify the variability in SRS planning practice between five experienced international centers. This planning study is based on 4 patient cases with representative tumor lesions in 1–2 vertebrae.

## Materials and methods

### Patient cases

All centers were experienced in spine RS planning and delivery and were members of the Elekta Spine Radiosurgery Research Consortium which included University Hospital Wuerzburg, University of Pittsburgh School of Medicine, William Beaumont Hospital, Princess Margaret Hospital, University of Virginia School of Medicine. Four cases from 4 institutions were selected that covered the range of spine RS practice. The epidural extension was Bilsky score 0, 1b and 2 and paraspinal involvement was observed in 3 of 4 cases [20]. Detailed information on patient and tumor characteristics is given in Table 1. Representative images from each case are illustrated in Fig. 1.

### Planning procedure

A step-wise planning procedure was performed such that variability in all stages of the planning process were analyzed independently: 1) CT-MR image registration 2) gross-tumor volume (GTV) and organ at risk (OAR) spinal cord delineation 3) clinical target volume (CTV) and planning risk volume (PRV) definition 4) planning target volume (PTV) definition 5) treatment planning based on consensus contours and 6) treatment planning based on institutional-specific contours. This was achieved by providing consensus results after each step of the planning process, forming the basis for the next step.

The planning software for treatment planning was Pinnacle<sup>3</sup> in all institutions (Philips Radiation Oncology

**Table 1** Detailed patient status parameters

Patient case:	1	2	3	4
Sex	F	F	F	M
Age	55	61	63	55
Histology:	Breast	Breast	Lung	Renal
Pain Score VAS	7	5	6	-
Paraspinal involvement:	Yes	No	Yes	Yes
Bilsky score:	1b	0	0	2
Circumferential epidural disease:	180°	90°	90°	180°
Location:	T 12-L1	C 2	T 7-8	L 3
Number of involved vertebrae:	2	1	2	1

Systems, Fitchburg, WI, USA) and some institutions used additional software for image registration or delineation depending on in house protocol.

### Step 1: CT-MR image registration

The CT and MR images for each patient were electronically provided to each institution. The CT images were defined as reference image data set and the MR image as secondary image set for image fusion. The registration parameters were recorded in x (Left-Right), y (Anterior-Posterior) and z (Superior-Inferior) directions. Only rigid translations and rotations of the image sets were considered for each case.

Average registration results were calculated by taking the mean value of all registration results and were considered to be the consensus results. These consensus registration parameters were sent back to all institutions. The consensus registration parameters were then used for the following step of target definition.

### Step 2: delineation of GTV and spinal cord

GTV and spinal cord (in one case with lumbar location - thecal sac) delineation was performed according to the institution specific in-house protocols in consideration of the clinical case descriptions by all institutions. A consensus for all case specific GTV and spinal cord contours was calculated and distributed to the different institutions by the coordinating institution.

### Step 3: delineation of CTV

Delineation of CTV was based on the consensus GTV. According to step 2, a consensus for all case specific contours was calculated and distributed to the different institutions by the coordinating institution.

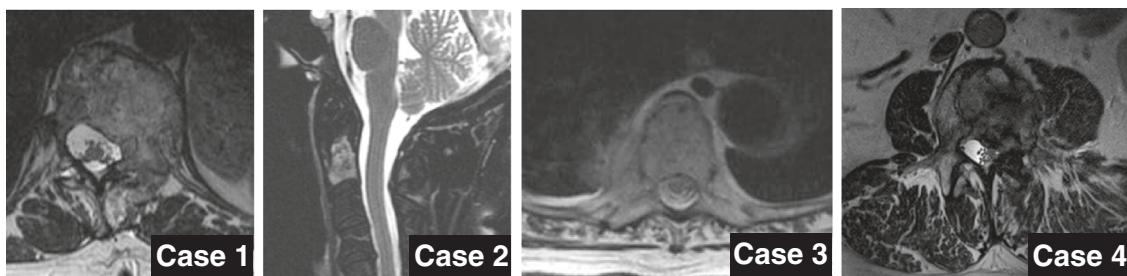
### Step 4: definition of PTV and PRV

PTV definition was based on the consensus CTV. The definition of the planning risk volume spinal cord (PRV\_SC) was based on the consensus spinal cord (in one case with lumbar location, consensus spinal canal). Once again, a consensus for all case specific contours was calculated and distributed to the different institutions by the coordinating institution.

### Step 5: treatment planning based on consensus contours

The consensus PTV and the consensus PRV\_SC were used for treatment planning. Institution specific contouring of relevant organs at risk (e.g. lungs...) and optimization guiding structures for the dose optimization process was allowed. Target volume prescription and normal tissue constraints are given below (see Step 6).

Institution specific in-house planning objectives/constraints were used for generation of the treatment plan. Both step-and-shoot IMRT as well as VMAT were allowed.



**Fig. 1** Representative slices from T2 weighted MR images of all patient cases. Case 1,3,4 are presented in axial view and case 2 in sagittal view

All institutions used an Elekta Synergy S/Axesee equipped with the Beam Modulator MLC (4 mm leaf width) for treatment planning (Elekta Beam Modulator™, Elekta Oncology Systems, Crawley, UK). Photon energies of 6 MV and 10 MV were used for treatment planning according to individual in-house protocols and depending on tumor location. Doses were calculated with a 2 mm grid size and collapsed cone convolution algorithm.

#### Step 6: treatment planning based on institutional-specific CTV contours

Treatment planning was repeated with institutional-specific GTV and CTV contours and institution specific PTV and OAR delineation by all institutions for all cases by using the identical planning objectives/constraints as in the previous planning step 5.

#### Objectives and constraints for this planning study

The prescription dose to the PTV was 18 Gy in a single fraction. An attempt was made to achieve a PTV coverage of at least 90 % by the prescription dose. Coverage of 80–90 % of the target volume was acceptable. A minimum coverage of <80 % of the target volume was an unacceptable deviation. Dose inhomogeneity within the target volume was allowed.

The dose to the normal tissue was limited by considering the tolerance dose values as listed in Table 2 which was adapted from RTOG 0631 Study [21].

#### Calculation of consensus structures

For each structure GTV, CTV and SC a consensus structure was determined such that each voxel encompassed by at least two institutions was included in the consensus structure. Only one outlier was excluded (Fig. 2).

#### Method for calculation of delineation variability

The consensus structure formed the basis for calculation of delineation variability. The shortest distance in each direction (Euclidean distance) was calculated between the consensus structure and the institutional structures

to obtain an estimation of delineation variability for these irregular shaped contours (Fig. 2).

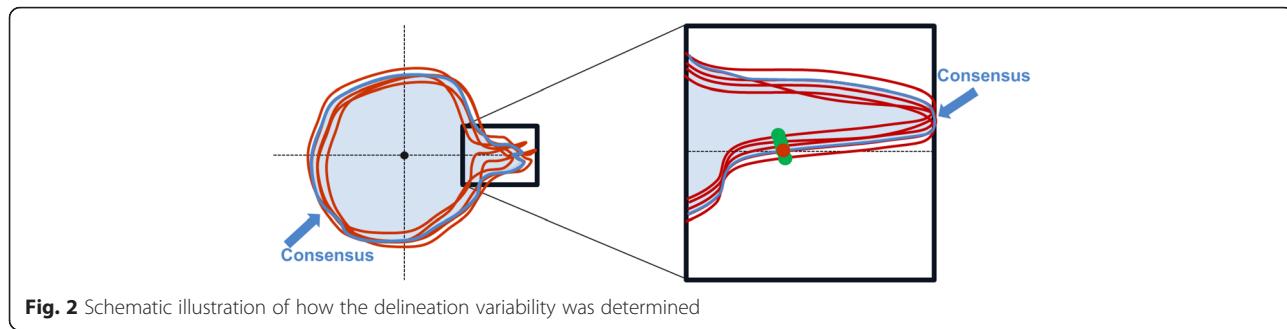
#### Method for dosimetric plan comparison

The dose distribution of the treatment plans was evaluated by determining minimum dose ( $D_{min}$ ), mean dose

**Table 2** Normal tissue constraints

Serial tissue	Volume in cm <sup>3</sup>	Dose Max in Gy	Endpoint (> Grade 3)
Spinal cord	<0.035	14	Myelitis
	<0.35	10	
	<1.2	7	
Cauda Equina	<0.035	16	Neuritis
	<5	14	
Sacral Plexus	<0.035	18	Neuropathy
	<5	14.4	
Esophagus <sup>a</sup>	<0.035	16	stenosis/fistula
	<5	11.9	
Heart/Pericardium	<0.035	22	Pericarditis
	<15	16	
Great vessels <sup>a</sup>	<0.035	37	Aneurysm
	<10	31	
Trachea <sup>a</sup> and Larynx	<0.035	20.2	stenosis/fistula
	<4	10.5	
Skin	<0.035	26	Ulceration
	<10	23	
Stomach	<0.035	16	ulceration/fistula
	<10	11.2	
Renal hilum/vascular trunk	<2/3 volume	10.6	malignant hypertension
Parallel tissue	Critical volume in cm <sup>3</sup>	Dose Max	Endpoint (> Grade 3)
Lung (Right & Left)	1000	7.4	Pneumonitis
Renal cortex (Right & Left)	200	8.4	Basic renal function

<sup>a</sup> Avoid circumferential irradiation



( $D_{\text{mean}}$ ), maximum dose ( $D_{\text{max}}$ ). The dose volume parameters according to ICRU Report 83 [22]  $D_{98}$ ,  $D_{95}$ ,  $D_{90}$ ,  $D_{05}$ ,  $D_{02}$  and  $V_{18\text{Gy}}$  were acquired for the PTV. The  $D_{98}$  of the PTV is the dose which encompasses 98 % of the PTV. Additionally,  $V_{18\text{Gy}}$  describes the part of the PTV which is treated with 18 Gy while  $V_{\text{Abs},18\text{Gy}}$  is the total body volume treated with 18 Gy (including normal tissue). For the PRV\_SC the parameters  $D_{\text{max}}$ ,  $D_{0,1\text{ccm}}$ ,  $V_{8\text{Gy}}$ ,  $V_{10\text{Gy}}$  were acquired.

The dose distributions were compared using the following parameters: conformity, homogeneity and target coverage, according to the RTOG recommendation [23]. Additionally, the Paddick conformity index was also determined [24]. The RTOG conformity index (CI-RTOG) represents the ratio of the volume encompassed by the prescription isodose to the target volume. Three categories of conformity index protocol compliance were defined. Plans with a conformity index value between 1.0 and 2.0 are in the normal range and are classified as not deviating from RTOG protocol. Conformity index value between 2.0 and 2.5 or between 0.9 and 1.0 are classified as having minor deviations. Value greater than 2.5 or less than 0.9 show major deviations of the dose from the RTOG protocol.

The RTOG homogeneity index (HI-RTOG) represents the ratio of maximum dose within the target volume to the prescribed dose. The normal range is up to 2.0, it indicates a minor deviation if  $\text{HI-RTOG} > 2.0$  and a major deviation if  $\text{HI-RTOG} > 2.5$ .

The RTOG coverage (Cov-RTOG) describes the ratio between minimum dose within the target volume and the prescribed dose, it indicates a minor deviation if  $<0.9$  and a major deviation if  $<0.8$ .

The Paddick conformity index (CI-Paddick) describes the ratio of the squared target volume covered by the prescription isodose to the arithmetic product of target volume and prescription isodose total volume. This equals the multiplication of the undertreatment ratio and the overtreatment ratio. This index has an ideal value of one and plan quality decreases with decreasing index value.

## Results

### Step 1: CT-MR image registration

The results of image registration are listed in Table 3. For each case, the standard deviation and maximum range between all institutional results were calculated.

**Table 3** Registration variability between the five institutions

	case	X in mm	Y in mm	Z in mm	rot X in °	rot Y in °	rot Z in °
Max range	<b>1</b>	4.7	2.2	2.7	3.1	0.8	4.4
	<b>2</b>	1.7	1.1	6.5	1.2	3.7	3.8
	<b>3</b>	1.8	15.9	11.8	4.1	2.9	4.5
	<b>4</b>	2.6	2.5	3.0	12.5	2.6	3.5
	<b>average 1–4</b>	2.7	5.4	6.0	5.2	2.5	4.1
	<b>average 1; 2; 4</b>	<b>3.0</b>	<b>1.9</b>	<b>4.1</b>	<b>5.6</b>	<b>2.4</b>	<b>3.9</b>
SD	<b>1</b>	1.7	0.9	1.0	1.1	0.4	1.6
	<b>2</b>	0.7	0.4	2.5	0.5	1.7	1.4
	<b>3</b>	0.7	6.9	4.8	1.7	1.2	1.8
	<b>4</b>	1.2	1.0	1.5	4.6	1.0	1.3
	<b>average 1–4</b>	1.1	2.3	2.5	2.0	1.1	1.5
	<b>average 1; 2; 4</b>	<b>1.2</b>	<b>0.8</b>	<b>1.7</b>	<b>2.1</b>	<b>1.0</b>	<b>1.4</b>

Abbreviations: x (left-right), y (anterior-posterior) and z (superior-inferior) direction; SD; standard deviation; average 1; 2; 4 (bold) excludes outlier case 3

Large registration variability was observed for case 3, most likely because of low image resolution (pixelspacing 1.2 mm CT and 0.8 mm on MRI) and large slice spacing of 3 mm on CT and 4.8 mm on MR. This was a thoracic spine (T 7–8) case with MR acquired in non-treatment position that made CT-MR registration extremely challenging. Without consideration of this outlier, average registration variability ranged from 0.8 mm to 1.7 mm for translation and 1.0° to 2.1° for rotation.

#### Step 2: delineation of GTV and spinal cord

The variation of the GTV contours is shown in Table 4. Variability was quantified by using the parameters SD and range in x (Left-Right), y (Anterior-Posterior) and z (Superior-Inferior) direction. On average the axial (X and Y) and longitudinal (Z) standard variation was 1.5 mm, 1.6 mm, respectively. The range was on average 3.5 mm and 3.8 mm in axial and longitudinal direction.

For the spinal cord, the delineation variability (1 SD) between the five institutions was 1 mm in axial direction averaged over the four cases.

#### Step 3: delineation of CTV and organ at risk spinal cord

Based on the consensus GTV, all institutions created a CTV which again resulted in variabilities of 0.8 mm and 1.2 mm in axial and longitudinal directions, respectively. The range was on average 1.8 mm and 2.8 mm in axial and longitudinal direction as shown in Table 5.

#### Step 4 delineation of PTV

CTV-to-PTV margins ranged between 0 mm and 2 mm according to institutional protocol [17, 25]. Margins for generation of the PRV spinal cord ranged between 1 mm and 2 mm.

#### Step 5: treatment planning

All participating institutions generated one treatment plan for each case using previously generated consensus structures. The dosimetric analyses are listed in Tables 6 and 7.

$D_{\min}$  to the PTV was  $7.5 \pm 1.8$  Gy averaged over all cases and institutions. The  $D_{90}$  was  $17.4 \pm 1.0$  Gy on average and 86.9 %  $\pm 5.2$  of the PTV was covered by the prescribed dose of 18 Gy.

**Table 4** GTV definition variability between the five institutions

Case	SD in mm		Range in mm	
	X and Y	Z	X and Y	Z
1	1.3	0.6	3.1	1.5
2	1.7	3.4	4.1	8.1
3	1.3	0.9	2.5	2.0
4	1.8	1.5	4.0	3.6
Average	1.5	1.6	3.5	3.8

**Table 5** CTV definition variability between the five institutions

Case	SD in mm		Range in mm	
	X and Y	Z	X and Y	Z
1	0.6	0.3	1.6	0.8
2	1.3	3.0	3.0	6.9
3	0.5	0.6	1.0	1.4
4	0.9	1.0	1.9	2.3
Average	0.8	1.2	1.8	2.8

$D_{\max}$  to the PRV\_SC<sub>consensus</sub> (spinal cord + 1 mm) was 10.5 Gy on average and variability (1 SD) of  $D_{\max}$  was 1.6 Gy averaged over all cases and institutions.  $D_{\max}$  in the PRV\_SC was smaller than 12 Gy in 18 of 20 treatment plan trials developed by the participating institutions. For the two deviating trials, the  $D_{0.1\text{ccm}}$  was 11 Gy in maximum. The highest variability between institutions was observed in case 4, where  $D_{\max}$  to the PRV\_SC ranged between 6.4 Gy and 11.7 Gy. The mean absolute PRV\_SC volume exposed to maximum 10 Gy was 0.1 cm<sup>3</sup> with a maximum of 0.39 cm<sup>3</sup>.

In Fig. 3 the ICRU dose report parameters are shown for one example case (case 1) for all 5 participating institutions. Good agreement is demonstrated for the parameters  $D_{90}$  and  $D_{\text{mean}}$  with a maximum deviation of 2.4 Gy and 1.1 Gy. A higher variability is demonstrated for the other dose report parameters – especially for  $D_{98}$ ,  $D_{95}$  and  $D_{\max}$  with a maximum difference of 4.9 Gy, 5.1 Gy and 4.6 Gy for this example case. Over all cases, the maximum deviation in  $D_{98}$  and  $D_{95}$  was 8.5 Gy and 7.6 Gy

Figure 4a illustrates the maximum dose to the planning risk volume spinal cord ( $D_{\max}$  PRV\_SC) as a function of the minimum dose in the PTV ( $D_{\min}$  PTV). Fig. 4b shows the dose to 0.1 cm<sup>3</sup> of the spinal cord to the  $D_{98}$  in the PTV. A very strong correlation can be seen - especially between  $D_{0.1\text{ccm}}(\text{SC})$  and  $D_{98}(\text{PTV})$  with a coefficient of determination of  $R^2 = 0.81$ .

For all cases and all plans, the performance parameters were calculated and analyzed. The detailed performance

**Table 6** ICRU report parameters with average values over all cases

	PTV	
$D_{\min}$ in Gy	7.5	$\pm 1.8$
$D_{98}$ in Gy	12.2	$\pm 2.0$
$D_{95}$ in Gy	15.2	$\pm 1.7$
$D_{90}$ in Gy	17.4	$\pm 1.0$
$D_{\text{mean}}$ in Gy	19.4	$\pm 0.8$
$D_{05}$ in Gy	21.6	$\pm 1.3$
$D_{02}$ in Gy	21.9	$\pm 1.4$
$D_{\max}$ in Gy	23.1	$\pm 1.6$

**Table 7** Mean of doses and volumes to PRV\_SC-Consensus

	D <sub>max</sub> in Gy	D <sub>0.1ccm</sub> in Gy	V <sub>8Gy</sub> in ccm	V <sub>10Gy</sub> in ccm	
Dose to PRV_SC <sub>consensus</sub> (Mean $\pm$ SD)	10.5 $\pm$ 1.6	9.0 $\pm$ 1.5	1.5 $\pm$ 1.6	0.1 $\pm$ 0.1	

parameters are listed in Table 8. Cov-RTOG was 87 % on average and variability was 5 %; variability was largest in case # 1 where PTV coverage ranged between 76 and 91 %. The CI-Paddick was 0.68 on average with small variability of 0.08; the largest variability in CI-Paddick was observed in case # 2 with a range of 0.76 and 0.91 between the institutions.

One additional plan was optimized based on the PTV<sub>individual</sub> of each case and institution. Doses to the consensus PTV and consensus PRV\_SC were evaluated in these treatment plans to quantify the dosimetric consequences of the image-registration and contouring uncertainties.

For the PTV<sub>individual</sub> based planning process, the variations on planning performance was higher compared to the consensus based planning. Performance parameters are illustrated in Figs. 5 and 6. All discrete values are illustrated with blue dots. Mean values over all institutions are shown in blue squares with standard deviation as error bars. In Fig. 5, the minor deviation range according to the respective protocol is shaded in light blue while major deviation rang is shown in shaded deeper blue.

Cov-RTOG decreased significantly if plans were generated based on institution-individual contours (1c - 4c). The mean values of consensus plans are centered in a range that would be considered as minor deviations ( $<0.9$ ) for all cases. This is contrasted with the mean of the individual plans (1i - 4i), which are consistently in a range that would be considered a major deviation ( $<0.8$ ) (Fig. 5).

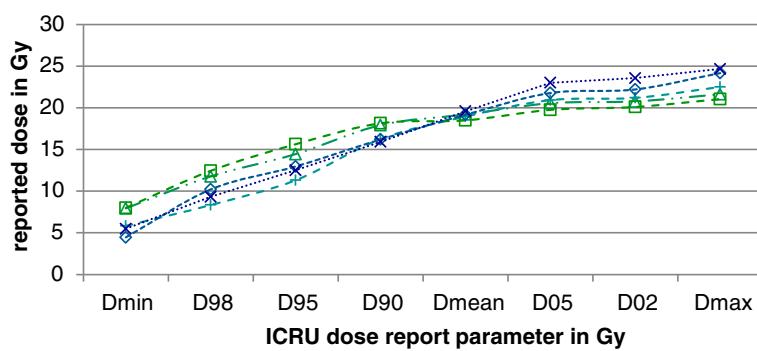
The CI-Paddick is less affected by the image registration and delineation variability. As shown in Fig. 6, the consensus plans show on average a higher CI-Paddick and smaller deviations.

For plans on individual PTVs, the average D<sub>min</sub> for PTV<sub>consensus</sub> drops from  $7.5 \pm 1.8$  Gy to  $6.5 \pm 0.9$  Gy. The D<sub>max</sub> to PRV\_SC increased from  $10.5 \pm 1.8$  Gy to  $12.2 \pm 2.2$  Gy. Similar results were observed for D<sub>0.1ccm</sub> which increased from D<sub>0.1ccm</sub>  $9.0 \pm 1.5$  Gy to  $9.9 \pm 1.8$  Gy.

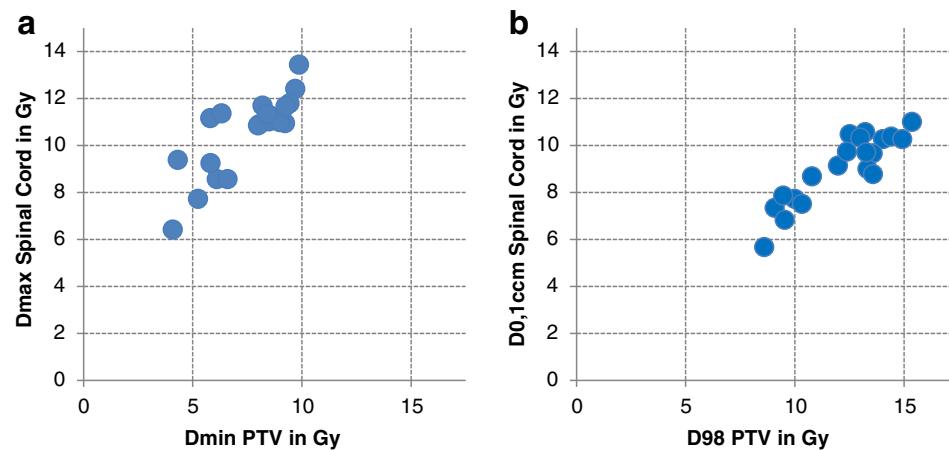
## Discussion

The aim of this work is to quantify the variability of treatment planning for spine radiosurgery in various steps of the treatment planning process between five international institutions. High variability was observed during all steps of the planning process of image registration and contouring of tumor and target volumes. For fair comparison, consensus registration parameters were determined for image fusion and consensus contours were developed for treatment plan optimization and dosimetric evaluation. Additional individual treatment plans were generated based on each individual institutions spine RS practice and showed higher variabilities in dose performance parameters for consensus target volumes and organs at risk.

For the CT-MRI registration process, a high variability was observed as reported by Ulin et al. [26]. None of the MR images were acquired in treatment position. The influence of the 3D- voxel size seems to be the main factor for registration performance. Case 1, 2 and 4 had an average voxel size for CT and MR datasets of 1.0 mm with an SD of 0.5 mm whereas case 3 had an voxel size of 4.1 mm for CT and 2.9 mm for MR dataset. Deformation of the vertebral column is well known but our results indicate that even in SBRT of a solitary vertebra, reproducible patient positioning for MR imaging should be performed to improve image registration. Additionally, all institutions explained that automatic image registration achieved unsatisfactory results and the image registration was adjusted manually. Improved software and methodologies specifically optimized for the vertebral column may therefore be required. One promising approach was suggested by Sohn et al. [27].



**Fig. 3** ICRU dose report parameters for one exemplary case (case 1) for all 5 participating institutions. The connecting lines should enhance the distinction between institutions



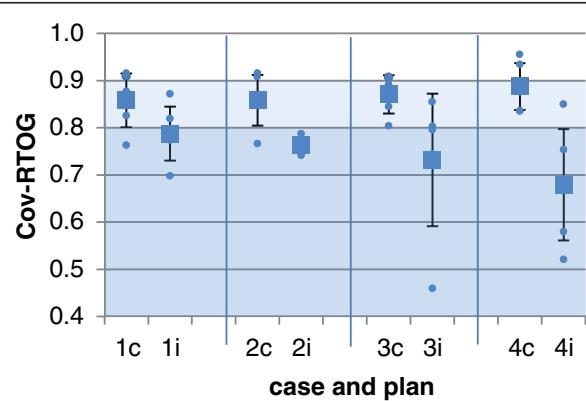
**Fig. 4 a/b** Maximum planning risk volume spinal cord (PRV\_SC) doses to PTV minimum doses and dose to 0,1ccm of spinal cord to PTV D98 for all cases and institutions

They present a segmental image fusion protocol which allows an improved visualization of spinal tumors and promises to achieve more consistent results. Nevertheless, image registration depends on datasets with high spatial resolution and a reproducible positioning to avoid or at least to minimize deformation.

For the delineation process, the highest variability was observed for GTV delineation, which was about 1.5 mm on average. The following step of CTV definition, which was based on the consensus GTV contour, resulted in lower variability between institutions of about 1 mm. The CTV-to-PTV margin ranged between 0 mm and 2 mm. These variabilities clearly show the need for further standardization of imaging and of delineation guidelines. Recently, GTV delineation variability was evaluated for SBRT of stage I non-small cell lung cancer and the overall delineation variability was 2.1 mm [28]. Tseng et al. results support the importance of controlling bulk patient motion and the practice of applying a planning organ-at-risk margin [29]. Several studies have investigated the inter-observer variability during the target definition of other entities. All of them agreed that the consistency of contouring can be improved by education and training, consensus guidelines and multi-institutional collaborations [30–33]. Also, a high degree of inter-observer variability was seen for brain SRS [34, 35]. Deviations between treatment planning based on individual and

consensus structures may show a potential influence on patient outcome.

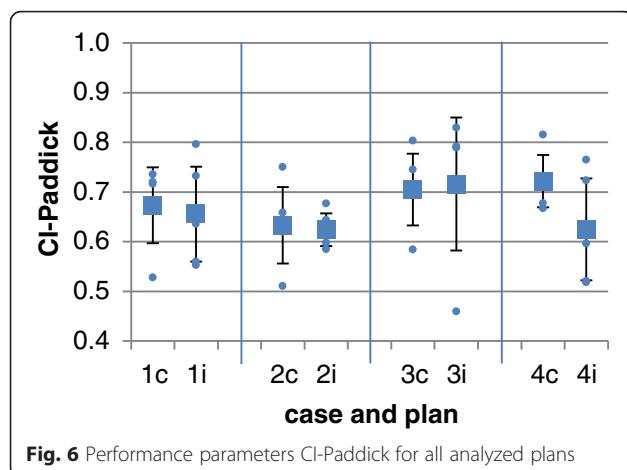
An overview of methods of spine radiosurgery for the participating institutions was described by Guckenberger et al. [17]. Good agreement was seen for the imaging acquisitioning techniques and safety margins concepts. However, treatment plan acceptance criteria varied substantially between all institutions. For this work, a  $D_{90}$  of 18 Gy in the PTV was suggested as parameter for plan acceptance. Even though the treatment planning was performed based on consensus contours, the acceptance criteria for  $D_{90}$  varied up to 3 Gy. This might be due to different approaches to achieve target coverage and sparing of the organs at risk, especially the spinal cord. The tolerance dose to the spinal cord is limited to a maximum dose while other institutions limit the dose to different sub-volumes ( $0.1 \text{ cm}^3$ ,  $0.4 \text{ cm}^3$  or  $0.04 \text{ cm}^3$ ). This disagreement was already stated by Guckenberger et al. [17]. Additionally, changes/deviations in beam configuration



**Fig. 5** Performance parameters RTOG coverage for all analyzed plans. Abbreviation: results of case 1–4 with consensus (c) and individual (i) plans

**Table 8** Mean performance parameters for all cases

Case	CI-Paddick	CI-RTOG	HI-RTOG	Cov-RTOG
1	0.67 $\pm 0.08$	1.11 $\pm 0.17$	1.98 $\pm 0.34$	0.86 $\pm 0.06$
2	0.63 $\pm 0.08$	1.19 $\pm 0.24$	1.64 $\pm 0.17$	0.86 $\pm 0.05$
3	0.70 $\pm 0.07$	1.09 $\pm 0.15$	1.90 $\pm 0.31$	0.87 $\pm 0.04$
4	0.72 $\pm 0.05$	1.10 $\pm 0.12$	1.90 $\pm 0.44$	0.89 $\pm 0.05$
mean	0.68 $\pm 0.08$	1.12 $\pm 0.18$	1.86 $\pm 0.35$	0.87 $\pm 0.05$



**Fig. 6** Performance parameters CI-Paddick for all analyzed plans

and optimization goals can lead to different planning results. Again, these variabilities show the need of strict and consistent acceptance criteria. Similar findings were published by Esposito et. al who investigated variability in treatment planning for stereotactic radiotherapy of liver metastasis [36]. In a multicenter study, they found significant differences for target coverage and OAR sparing due to different optimization strategies selected by the planners. Another multi-institutional study evaluated dosimetric parameters for SBRT of lung lesions and how much the guidelines provided in the literature are being successfully implemented in a variety of clinics [37]. They analyzed PTV coverage and conformality of their treatment plan and found that the conformality index of 50 % was the most difficult to meet depending on tumor size and location.

## Conclusion

Spinal radiosurgery (SRS) for vertebral metastases is a rapidly evolving treatment modality that has shown promising results in terms of pain and tumor control. However, the methodology and implementation of SRS has not yet been sufficiently standardized. Further studies are needed to establish whether the variability observed in this study will influence the clinical outcomes.

## Competing interests

This research was partially supported through an Elektro research grant with all institutions being members of the Elektro Spine Radiosurgery Research Consortium. Dr. Arjun Sahgal has received honorarium for past educational seminars from Elektro AB and research grants from Elektro. Dr. Grills owns stock in and is on the board of directors of Greater Michigan Gamma Knife.

## Authors' contributions

MG designed the study and MG, AT and AR performed the data analysis. All authors participated in data collection and performed critical review of the manuscript and finally approved the manuscript.

## Acknowledgements

This research was partially supported through an Elektro research grant with all institutions being members of the Elektro Spine Radiosurgery Research Consortium. This publication was supported by the Open Access Publication

Fund of the University of Wuerzburg. This work and these data, however, are the intellectual property of the individual group members and their sponsoring institutions.

## Author details

<sup>1</sup>Department of Radiation Oncology, University of Wuerzburg, Wuerzburg, Germany. <sup>2</sup>Departments of Neurological Surgery and Radiation Oncology, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA. <sup>3</sup>William Beaumont Hospital, Royal Oak, MI, USA. <sup>4</sup>Department of Radiation Oncology, Sunnybrook Odette Cancer Center, University of Toronto, Toronto, ON, Canada. <sup>5</sup>Princess Margaret Cancer Center, Toronto, ON, Canada. <sup>6</sup>University of Virginia School of Medicine, Charlottesville, VA, USA. <sup>7</sup>Division of Radiation Oncology, University Hospital Zurich, Zurich, Switzerland.

Received: 15 March 2016 Accepted: 9 April 2016

Published online: 18 April 2016

## References

- Witham TF, Khavkin YA, Gallia GL, Wolinsky JP, Gokaslan ZL. Surgery insight: current management of epidural spinal cord compression from metastatic spine disease. *Nat Clin Pract Neurol*. 2006;2:87–94. quiz 116.
- Sundaresan N, Borhani S, Rothman A, Holtzman R. Tumors of the osseous spine. *J Neurooncol*. 2004;69:273–90.
- Wong DA, Fornasier VL, MacNab I. Spinal metastases: the obvious, the occult, and the impostors. *Spine (Phila Pa 1976)*. 1990;15:1–4.
- De Bari B, Alongi F, Mortellaro G, Mazzola R, Schiappacasse L, Guckenberger M. Spinal metastases: Is stereotactic body radiation therapy supported by evidences? *Crit Rev Oncol Hematol*. 2016;98:147–58. doi:10.1016/j.critrevonc.2015.11.002. Epub 2015 Nov 10.
- Lutz S, Berk L, Chang E, Chow E, Hahn C, Hoskin P, Howell D, Konski A, Kachnic L, Lo S, et al. Palliative radiotherapy for bone metastases: an ASTRO evidence-based guideline. *Int J Radiat Oncol Biol Phys*. 2011;79: 965–76.
- Souchon R, Wenz F, Sedlmayer F, Budach W, Dunst J, Feyer P, Haase W, Harms W, Sautter-Bihl ML, Sauer R, German Society of Radiation O. DEGRO practice guidelines for palliative radiotherapy of metastatic breast cancer: bone metastases and metastatic spinal cord compression (MSCC). *Strahlenther Onkol*. 2009;185:417–24.
- Chow E, Harris K, Fan G, Tsao M, Sze WM. Palliative radiotherapy trials for bone metastases: a systematic review. *J Clin Oncol*. 2007;25:1423–36.
- Mizumoto M, Harada H, Asakura H, Hashimoto T, Furutani K, Hashii H, Takagi T, Katagiri H, Takahashi M, Nishimura T. Prognostic factors and a scoring system for survival after radiotherapy for metastases to the spinal column: a review of 544 patients at Shizuoka Cancer Center Hospital. *Cancer*. 2008;113: 2816–22.
- Leithner A, Radl R, Gruber G, Hochegger M, Leithner K, Welkerling H, Rehak P, Windhager R. Predictive value of seven preoperative prognostic scoring systems for spinal metastases. *Eur Spine J*. 2008;17:1488–95.
- van der Linden YM, Dijkstra SP, Vonk EJ, Marijnen CA, Leer JW, Dutch Bone Metastasis Study G. Prediction of survival in patients with metastases in the spinal column: results based on a randomized trial of radiotherapy. *Cancer*. 2005;103:320–8.
- Guckenberger M, Mantel F, Gerszten PC, Flickinger JC, Sahgal A, Letourneau D, Grills IS, Jawad M, Fahim DK, Shin JH, et al. Safety and efficacy of stereotactic body radiotherapy as primary treatment for vertebral metastases: a multi-institutional analysis. *Radiat Oncol*. 2014;9:226.
- Mantel F, Glatz S, Toussaint A, Flentje M, Guckenberger M. Long-term safety and efficacy of fractionated stereotactic body radiation therapy for spinal metastases. *Strahlenther Onkol*. 2014;190:1141–8.
- Laufer I, Iorgulescu JB, Chapman T, Lis E, Shi W, Zhang Z, Cox BW, Yamada Y, Bilsky MH. Local disease control for spinal metastases following "separation surgery" and adjuvant hypofractionated or high-dose single-fraction stereotactic radiosurgery: outcome analysis in 186 patients. *J Neurosurg Spine*. 2013;18:207–14.
- Schipani S, Wen W, Jin JY, Kim JK, Ryu S. Spine radiosurgery: a dosimetric analysis in 124 patients who received 18 Gy. *Int J Radiat Oncol Biol Phys*. 2012;84:e571–6.
- Nguyen QN, Shiu AS, Rhines LD, Wang H, Allen PK, Wang XS, Chang EL. Management of spinal metastases from renal cell carcinoma using stereotactic body radiotherapy. *Int J Radiat Oncol Biol Phys*. 2010;76:1185–92.

16. Gerszten PC, Burton SA, Ozhasoglu C, Welch WC. Radiosurgery for spinal metastases: clinical experience in 500 cases from a single institution. *Spine*. 2007;32:193–9.
17. Guckenberger M, Sweeney RA, Flickinger JC, Gerszten PC, Kersh R, Sheehan J, et al. Clinical practice of image-guided spine radiosurgery—results from an international research consortium. *Radiat Oncol*. 2011;6:172.
18. Gerszten PC, Sahgal A, Sheehan JP, Kersh R, Chen S, Flickinger JC, Quader M, Fahim D, Grills I, Shin JH, et al. A multi-national report on methods for institutional credentialing for spine radiosurgery. *Radiat Oncol*. 2013;8:158.
19. Dahele M, Hatton M, Slotman B, Guckenberger M. Stereotactic body radiotherapy: a survey of contemporary practice in six selected European countries. *Acta Oncol*. 2015;54:1237–41.
20. Bilsky MH, Laufer I, Fourney DR, Groff M, Schmidt MH, Varga PP, Vrionis FD, Yamada Y, Gerszten PC, Kuklo TR. Reliability analysis of the epidural spinal cord compression scale. *J Neurosurg Spine*. 2010;13:324–8.
21. Ryu S, Pugh SL, Gerszten PC, Yin FF, Timmerman RD, Hitchcock YJ, Movsas B, Kanner AA, Berk LB, Followill DS, Kachnic LA. RTOG 0631 phase II/III study of image-guided stereotactic radiosurgery for localized (1–3) spine metastases: phase II results. *Int J Radiat Oncol Biol Phys*. 2011;81:S131–2.
22. Prescribing, recording, and reporting photon-beam Intensity-Modulated Radiation Therapy (IMRT): contents. *J ICRU* 2010, 10:NP.
23. Shaw E, Kline R, Gillin M, Souhami L, Hirschfeld A, Dinapoli R, Martin L. Radiation Therapy Oncology Group: radiosurgery quality assurance guidelines. *Int J Radiat Oncol Biol Phys*. 1993;27:1231–9.
24. Paddick I. A simple scoring ratio to index the conformity of radiosurgical treatment plans. Technical note. *J Neurosurg*. 2000;93 Suppl 3:219–22.
25. Li W, Sahgal A, Foote M, Millar BA, Jaffray DA, Letourneau D. Impact of immobilization on intrafraction motion for spine stereotactic body radiotherapy using cone beam computed tomography. *Int J Radiat Oncol Biol Phys*. 2012;84:520–6.
26. Ulin K, Urie MM, Cherlow JM. Results of a multi-institutional benchmark test for cranial CT/MR image registration. *Int J Radiat Oncol Biol Phys*. 2010;77: 1584–9.
27. Sohn MJ, Lee DJ, Yoon SW, Lee HR, Hwang YJ. The effective application of segmental image fusion in spinal radiosurgery for improved targeting of spinal tumours. *Acta Neurochir (Wien)*. 2009;151:231–8. discussion 238.
28. Peulen H, Belderbos J, Guckenberger M, Hope A, Grill I, van Herk M, Sonke JJ. Target delineation variability and corresponding margins of peripheral early stage NSCLC treated with stereotactic body radiotherapy. *Radiat Oncol*. 2015;11:361–6.
29. Tseng CL, Sussman MS, Atenafu EG, Letourneau D, Ma L, Soliman H, Thibault I, Cho BC, Simeonov A, Yu E, et al. Magnetic resonance imaging assessment of spinal cord and cauda equina motion in supine patients with spinal metastases planned for spine stereotactic body radiation therapy. *Int J Radiat Oncol Biol Phys*. 2015;91:995–1002.
30. Schimek-Jasch T, Troost EG, Rucker G, Prokic V, Avlar M, Duncker-Rohr V, Mix M, Doll C, Grosu AL, Nestle U. A teaching intervention in a contouring dummy run improved target volume delineation in locally advanced non-small cell lung cancer: reducing the interobserver variability in multicentre clinical studies. *Strahlenther Onkol*. 2015;191:525–33.
31. Eminowicz G, McCormack M. Variability of clinical target volume delineation for definitive radiotherapy in cervix cancer. *Radiat Oncol*. 2015.
32. Brouwer CL, Steenbakkers RJ, Bourhis J, Budach W, Grau C, Gregoire V, van Herk M, Lee A, Maingon P, Nutting C, et al. CT-based delineation of organs at risk in the head and neck region: DAHANCA, EORTC, GORTEC, HKNPCSG, NCIC CTG, NCRI, NRG Oncology and TROG consensus guidelines. *Radiat Oncol*. 2015;11:83–90.
33. Major T, Gutierrez C, Guix B, Moza E, Hannoun-Levi JM, Lossi K, Niehoff P, Resch A, van Limbergen E, Polgar C. Interobserver variations of target volume delineation in multicatheter partial breast brachytherapy after open cavity surgery. *Brachytherapy*. 2015.
34. Sandstrom H, Nordstrom H, Johansson J, Kjall P, Jokura H, Toma-Dasu I. Variability in target delineation for cavernous sinus meningioma and anaplastic astrocytoma in stereotactic radiosurgery with Leksell Gamma Knife Perfexion. *Acta Neurochir (Wien)*. 2014;156:2303–12. discussion 2312–2303.
35. Stanley J, Dunscombe P, Lau H, Burns P, Lim G, Liu HW, Nordal R, Starreveld Y, Valev B, Voroney JP, Spencer DP. The effect of contouring variability on dosimetric parameters for brain metastases treated with stereotactic radiosurgery. *Int J Radiat Oncol Biol Phys*. 2013;87:924–31.
36. Esposito M, Maggi G, Marino C, Bottalico L, Cagni E, Carbonini C, Casale M, Clemente S, D'Alesio V, Fedele D, et al. Multicentre treatment planning inter-comparison in a national context: the liver stereotactic ablative radiotherapy case. *Phys Med*. 2015.
37. Woerner A, Roeske JC, Harkenrider MM, Fan J, Aydogan B, Koshy M, Laureckas R, Vali F, Campana M, Surucu M. A multi-institutional study to assess adherence to lung stereotactic body radiotherapy planning goals. *Med Phys*. 2015;42:4629–35.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

