International Journal of Radiation Oncology biology • physics

www.redjournal.org

Clinical Investigation: Central Nervous System Tumor

International Spine Radiosurgery Consortium Consensus Guidelines for Target Volume Definition in Spinal Stereotactic Radiosurgery

Brett W. Cox, MD,[*,1] Daniel E. Spratt, MD,[*,1] Michael Lovelock, PhD,[†] Mark H. Bilsky, MD,[‡] Eric Lis, MD,[§] Samuel Ryu, MD,[‖] Jason Sheehan, MD,[¶] Peter C. Gerszten, MD, MPH,[**] Eric Chang, MD,[††] Iris Gibbs, MD,[‡‡] Scott Soltys, MD,[‡‡] Arjun Sahgal, MD,[§§] Joe Deasy, PhD,[†] John Flickinger, MD,[‖‖] Mubina Quader, PhD,[‖‖] Stefan Mindea, MD,[¶¶] and Yoshiya Yamada, MD[‡‡]

*Departments of *Radiation Oncology, [†]Medical Physics, [‡]Surgery, Division of Neurosurgery, and [§]Radiology, Memorial Sloan-Kettering Cancer Center, New York, New York; [‖]Department of Radiation Oncology, Henry Ford Hospital, Detroit, Michigan; [¶]Department of Neurosurgery, University of Virginia School of Medicine, Charlottesville, Virginia; **Department of Neurological Surgery and Radiation Oncology, University of Pittsburgh Medical Center, UPMC Presbyterian, Pittsburgh, Pennsylvania; [††]Department of Radiation Oncology, University of Southern California Keck School of Medicine, Health Sciences Campus, Los Angeles, California; [‡‡]Department of Radiation Oncology, Stanford University School of Medicine, Stanford, California; [§§]Department of Radiation Oncology, Princess Margaret Hospital and the Sunnybrook Health Sciences Center, University of Toronto, Toronto, Ontario, Canada; [‖‖]Department of Radiation Oncology, University of Pittsburgh Medical Center, UPMC Presbyterian, Pittsburgh, Pennsylvania; and [¶¶]Department of Neurosurgery, Stanford University School of Medicine, Stanford, California*

Received Dec 21, 2011, and in revised form Feb 24, 2012. Accepted for publication Mar 2, 2012

Summary

Ten physician members of the International Spine Radiosurgery Consortium independently contoured 10 cases representing common scenarios in spinal radiosurgery for metastases. Target volumes were

Purpose: Spinal stereotactic radiosurgery (SRS) is increasingly used to manage spinal metastases. However, target volume definition varies considerably and no consensus target volume guidelines exist. This study proposes consensus target volume definitions using common scenarios in metastatic spine radiosurgery.

Methods and Materials: Seven radiation oncologists and 3 neurological surgeons with spinal radiosurgery expertise independently contoured target and critical normal structures for 10 cases representing common scenarios in metastatic spine radiosurgery. Each set of volumes was imported into the Computational Environment for Radiotherapy Research. Quantitative analysis was performed using an expectation maximization algorithm for Simultaneous Truth and Performance Level Estimation (STAPLE) with kappa statistics calculating agreement

Note—An online CME test for this article can be taken at http://astro.org/MOC.

Reprint requests to: Brett W. Cox, MD, Department of Radiation Oncology, Memorial Sloan-Kettering Cancer Center, 1275 York Ave, New York, NY 10065. Tel: (212) 639-5717; Fax: (212) 639-2417; E-mail: coxb@mskcc.org

M.H.B. receives royalties from Stackable Cage, Depuy Spine. P.C.G. received payment for lectures and expenses from Elekta, AB. I.G. was a member of a Clinical Advisory Board for Accuray, Inc, and received payment for lectures from Accuray, Inc Y.Y. is a consultant for Varian Medical Systems and Funcacion Hospital Provincial Castellon. E.C., B.W.C., J.D., J.F., E.L., M.L., S.M., M.Q., S.R., A.S., J.S., S.S., and D.E.S. have no disclosures to report.

[1] Coauthors.

Int J Radiation Oncol Biol Phys, Vol. 83, No. 5, pp. e597–e605, 2012

doi:10.1016/j.ijrobp.2012.03.009

quantitatively analyzed using an expectation maximization algorithm for Simultaneous Truth and Performance Level Estimation. Consensus guidelines for target volume definition in spinal stereotactic radiosurgery for metastatic disease were generated. This report serves as a foundation for refining radiosurgery target volume delineation. We advocate using consensus target definitions in future spine radiosurgery protocols.

between physicians. Optimized confidence level consensus contours were identified using histogram agreement analysis and characterized to create target volume definition guidelines.
Results: Mean STAPLE agreement sensitivity and specificity was 0.76 (range, 0.67-0.84) and 0.97 (range, 0.94-0.99), respectively, for gross tumor volume (GTV) and 0.79 (range, 0.66-0.91) and 0.96 (range, 0.92-0.98), respectively, for clinical target volume (CTV). Mean kappa agreement was 0.65 (range, 0.54-0.79) for GTV and 0.64 (range, 0.54-0.82) for CTV ($P<.01$ for GTV and CTV in all cases). STAPLE histogram agreement analysis identified optimal consensus contours (80% confidence limit). Consensus recommendations include that the CTV should include abnormal marrow signal suspicious for microscopic invasion and an adjacent normal bony expansion to account for subclinical tumor spread in the marrow space. No epidural CTV expansion is recommended without epidural disease, and circumferential CTVs encircling the cord should be used only when the vertebral body, bilateral pedicles/lamina, and spinous process are all involved or there is extensive metastatic disease along the circumference of the epidural space.
Conclusions: This report provides consensus guidelines for target volume definition for spinal metastases receiving upfront SRS in common clinical situations.

Introduction

Spinal stereotactic radiosurgery (SRS), also known as spinal stereotactic body radiation therapy, is being increasingly used in the management of metastatic spine tumors (1-3). SRS uses a combination of image-guided technologies to remediate inter- and intra-fraction motion and advanced inverse treatment-planning algorithms to achieve highly conformal dose distributions (4). This permits delivery of treatment plans with steep dose gradients between target volumes and adjacent organs at risk (5, 6). This is particularly desirable for treating spinal tumors, where local failure is catastrophic and tumors abut sensitive critical normal tissues such as the spinal cord and esophagus (7, 8). Approximately 50% of radiation oncologists currently incorporate spinal SRS into their practice (9), and there is widespread acceptance of SRS in the neurosurgical community (1, 10, 11).

Large retrospective single-institution studies demonstrate the efficacy of spinal SRS with durable local control, progression-free survival, and palliation rates of approximately 70%-90% in patients receiving *de novo* treatment, retreatment after prior radiation, and treatment postoperatively after surgical resection (12-14). Radiosurgery has also shown potential for decompressing spine metastases with epidural extension causing spinal cord compression (15). The results of 2 prospective studies for spinal radiosurgery have recently been reported (16, 17), both concluding that spinal radiosurgery is feasible, effective, and indicated as an expanding first-line option in properly selected patients with spinal metastases. Radiation Therapy Oncology Group 0631 is ongoing and is accruing to the phase III component, which compares up to 18 Gy of spinal SRS with a control arm of 8 Gy, both delivered in a single fraction.

Although there are limited published guidelines regarding spine SRS indications and patterns of practice (7), consensus definitions for target volumes have yet to be established (3, 18). Consensus definitions are necessary to standardize the nomenclature and delivery of spinal radiosurgery and enable comparison of results from different institutions, treatment platforms, and dose-fraction schedules. Patterns of failure analysis have identified marginal failures beyond the conformal targeted volume as a primary pattern of recurrence (19, 20), highlighting the need for accuracy in target volume delineation. Identifying the correct clinical tumor volume is of particular importance because the steep dose gradients associated with SRS result in subtherapeutic doses within millimeters of the planning target volume, and the adjacent normal tissues are at risk of injury from high dose-per-fraction regimens.

The purpose of this report is to provide consensus guidelines for target volume definition for spinal metastases when using SRS.

Methods and Materials

The International Spine Radiosurgery Consortium (ISRC) consists of radiosurgery experts from the Memorial Sloan-Kettering Cancer Center, Henry Ford Hospital, University of Virginia School of Medicine, University of Pittsburgh Medical Center, University of Southern California Keck School of Medicine, Stanford University School of Medicine, University of Toronto (Princess Margaret Hospital and Sunnybrook Health Sciences Center), and Cleveland Clinic. The ISRC was organized to address important issues surrounding contemporary spine radiosurgery. One of the earliest initiatives of the consortium was to propose standardization of contouring guidelines for spinal SRS using formal International Commission on Radiation Units and Measurements nomenclature for definition of target volumes.

Seven radiation oncologists and 3 neurological surgeons with extensive expertise and a collective experience of more than 2000 spine SRS cases agreed to serve on a task force assigned to create consensus target volume definitions for SRS for spinal metastases. Ten cases from actual patients treated with SRS were identified by the task force to represent commonly encountered situations in spinal SRS. Cases were specifically selected to ascertain how the metastatic lesion's anatomic location within the vertebral structure and the degree of any epidural or paraspinal extension impacted target volume delineation.

The contouring physician was provided with a clinical vignette including patient age, oncologic history, pertinent signs and symptoms, and relevant radiology report findings. A complete set of pre- and postcontrast multiplanar magnetic resonance imaging (MRI) was provided for review in DICOM format. Panel members were instructed to assume all cases were to be treated with SRS. The panel members independently contoured the gross target volume (GTV), clinical target volume (CTV), planning target

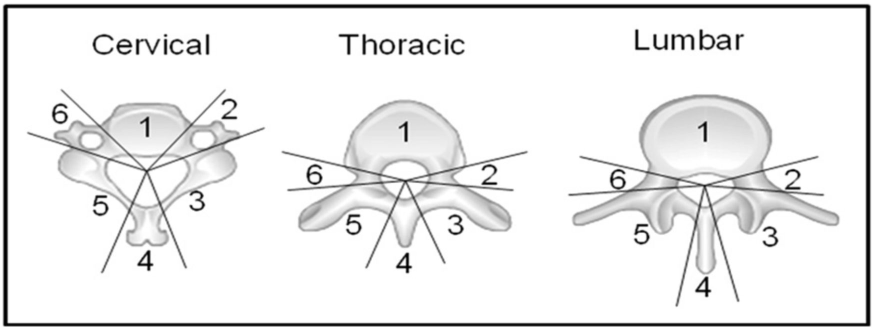


Fig. 1. International Spine Radiosurgery Consortium anatomic classification system for consensus target volumes for spine radiosurgery.

volume (PTV), and spinal cord and/or cauda equina on provided DICOM-format computer tomography (CT) treatment-planning simulation scans. All patients were simulated supine in a custom cradle with a myelogram used in selected cases to assist with delineation of the spinal cord and any epidural disease. Axial CT images were obtained at 2-mm intervals. MRI fusion was permitted for treatment-planning purposes. Physicians also answered a short survey regarding contour delineation rationale for each case.

The completed volumes were centrally collected in DICOM format and placed into in-house treatment-planning software in which all contours for each case were co-registered and superimposed on a single CT scan. Each case was then imported into the Computational Environment for Radiotherapy Research program (CERR Version 4.0) run through MATLAB (Version 7.12.0) for analysis. It was assumed that the true target volumes were contained within the overlaid target volumes. Quantitative analysis was performed using an expectation maximization algorithm for Simultaneous Truth and Performance Level Estimation (STAPLE) to quantify expert agreement (21). This algorithm calculates a probabilistic estimate of a true segmentation from an input collection of segmentations, allowing for global analysis and quantification of the performance level of each collective segmentation. Allozi et al provide further details regarding the specific algorithms and statistical methods of the STAPLE and kappa metrics used (22). STAPLE analysis enables a systematic approach to developing consensus contours, with sensitivity, specificity, and agreement level measurements helping to identify the best clinical agreement consensus contour. To account for the possibility of expert agreement by chance, kappa statistics were used to represent agreement between contouring physicians. A kappa statistic of 1 represents perfect agreement between experts; a value of 0 represents no agreement above that which might have be seen through random chance (23). Based on STAPLE analysis from the overlaid contours from each case, consensus contours for the GTV and CTV were generated using the 80% confidence level. All GTV consensus contours were reviewed by a neuroradiologist to ensure accuracy. CTV consensus contours were characterized. Recommendations regarding target volume definition were generated.

The task force recommended that anatomic descriptions of consensus contours be made using a modified Weinstein-Boriani-Biagini system for spine tumor staging (24). This is a standardized staging system that has been widely adopted in the neurological and orthopedic surgical literature to facilitate surgical planning and communication of results from *en bloc* vertebrectomy. It was the opinion of the task force that dividing each vertebral body into 12 sectors, as done in the Weinstein-Boriani-Biagini system, would be unwieldy for consensus contour recommendations. Therefore, the modified system proposed here divides each vertebral body into 6 sectors, as shown in Fig. 1. Sector 1 represents the vertebral body, sector 2 represents the left pedicle, sector 3 represents the left transverse process and lamina, sector 4 represents the spinous process, sector 5 represents the right transverse process and lamina, and sector 6 represents the right pedicle.

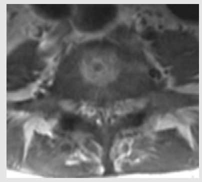
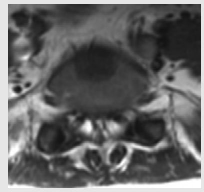
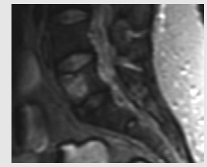
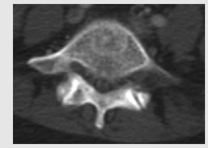
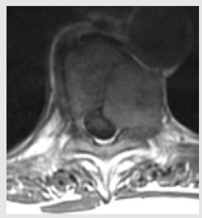
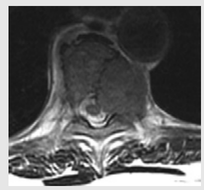
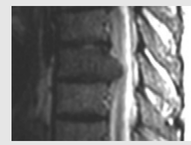
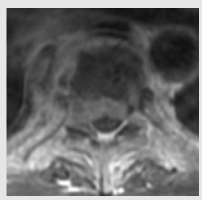
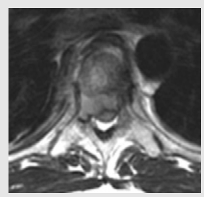
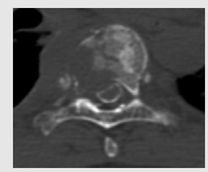
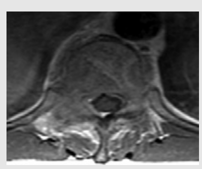
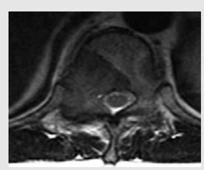
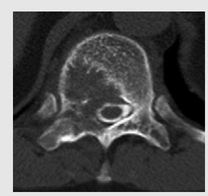
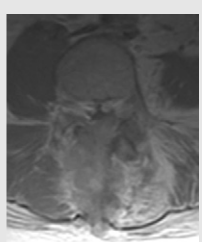
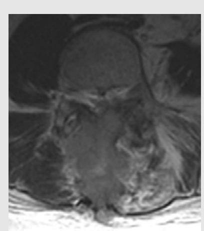
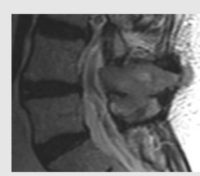
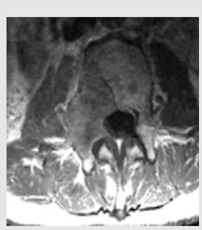
Results

Case cohort characteristics are detailed in Table 1, with anatomic descriptions and representative MRI T1-weighted axial post-contrast, MRI T2-weighted axial, MRI T2 sagittal, and axial CT imaging available for each case. Six cases involved the thoracic spine and 4 cases involved the lumbar spine. Seven cases had a component of epidural disease, with 3 of these cases demonstrating deformation of the thecal sac. During the analysis, approximately 400 total contoured structures (target volumes and organs at risk), distributed over more than 7500 total axial CT slices, were analyzed. Nine of 10 task force members completed a GTV, CTV, and PTV for each case. The GTV and CTV contour completion rate was 100%. Only completed structures were included in the analysis. There were no apparent variations in GTV or CTV target contour definition between the radiation oncologists and neurosurgeons.

GTV delineation

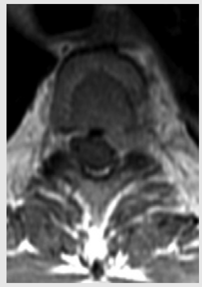
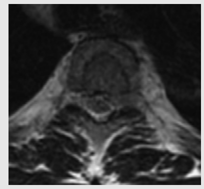
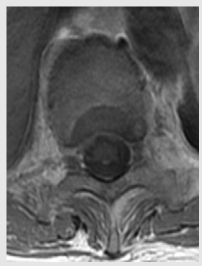
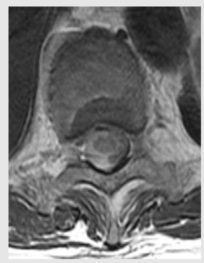
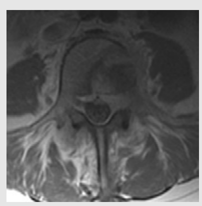
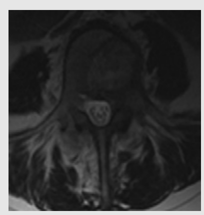
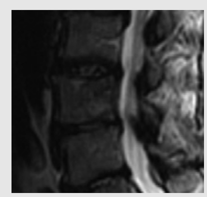
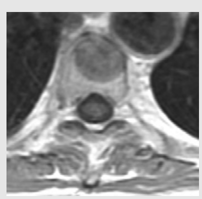
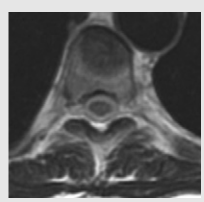
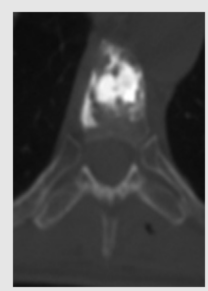
For the GTV contours in all 10 cases, there was a substantial level of STAPLE agreement between contouring physicians with mean sensitivity and specificity rates for each case of 0.76 (range, 0.67-0.84) and 0.97 (range, 0.94-0.99), respectively (Table 2). The mean GTV kappa agreement level was 0.65 (range, 0.54-0.79). The kappa agreement level of the GTV was statistically significant with $P<.01$ for all 10 cases. Agreement histograms for all cases demonstrated plateaued GTV volume agreement from the 0.1-0.9 confidence levels. Apparent and corrected (kappa) agreement histograms revealed decreasing absolute volume agreement with increasing confidence levels from 0-1. Review of all GTV 80% confidence level consensus contours by an expert neuroradiologist confirmed correct and complete delineation of the gross tumor including all bony, epidural, and paraspinal components.

Table 1 Summary of cases for spinal SRS consensus contours

Anatomic description	MRI axial T1 post	MRI axial T2	MRI sagittal	CT
Case 1: L5 lesion limited to the anterior VB with no epidural extension				
Case 2: T8 lesion involving left pedicle, posterolateral VB, and neural foramen. Involvement of the ventral and left lateral epidural space, mild spinal canal compromise, and abutment of the spinal cord				
Case 3: T6-8 lesion with T6 collapse deformity, ventral epidural disease, moderate spinal canal compromise, mild spinal cord displacement, extension to the bilateral neural foramina, and paraspinal extension				
Case 4: T11 lesion involving pedicle and posterior elements, mild ventral and right lateral epidural disease, narrowing of the right T10/11 and T11/12 neural foramina				
Case 5: L5 lesion centered in the spinous process and extending to the bilateral lamina, bilateral posterior paraspinal musculature, and bilateral dorsal epidural space extension with mild spinal canal compromise				
Case 6: L2-3 expansile mass in right-sided VB and right posterior elements with mild right ventral, lateral, and dorsal epidural disease. Involvement of the right L2/3 and L3/4 neural foramina				

(*continued on next page*)

Table 1 *(continued)*

Anatomic description	MRI axial T1 post	MRI axial T2	MRI sagittal	CT
Case 7: T3 posterior VB lesion extending into the left neural foramen with mild spinal canal compromise and left ventral and lateral epidural extension				
Case 8: T10 lesion in posterior VB				
Case 9: L4 diffuse marrow replacement including left pedicle and articular facets, ventral epidural extension, left lateral recess extension, and left L4/5 neural foramen involvement				
Case 10: T5 lesion with mild superior and inferior endplate infractions resulting in mild loss of VB height. Mild anterior paraspinal extension. Patient underwent T5 kyphoplasty				

Abbreviations: CT = computed tomography; MRI = magnetic resonance imaging; VB = vertebral body.

CTV delineation

STAPLE analysis of the CTV (Table 2) also revealed a high level of agreement, with sensitivity and specificity rates of 0.79 (range, 0.66-0.91) and 0.96 (range, 0.92-0.98). The mean kappa agreement level was 0.64 (range, 0.54-0.82) with the kappa agreement $P<.01$ for all 10 CTV contours. STAPLE agreement histograms demonstrate stable CTV volume agreement from the 0.1-0.9 confidence levels for all cases but decreasing volume agreement through the same range of confidence levels using the apparent and corrected algorithms. Figure 2 demonstrates individual CTV contours (shown in red) and overlaid 80% consensus contours (shown in orange) as well as a schematic of CTV contour recommendations based on the ISRC anatomic classification system. Axial MRI images are again provided for reference.

Consensus GTV and CTV definitions

An 80% confidence level threshold was selected for the GTV and CTV consensus contours based on analysis of the STAPLE agreement histograms to maximize overall STAPLE absolute

Table 2 GTV and CTV contour agreement between consortium members

	GTV				CTV			
	Mean sensitivity (avg. ± SD)	Mean specificity (avg. ± SD)	Kappa measure	*P* value	Mean sensitivity (avg. ± SD)	Mean specificity (avg. ± SD)	Kappa measure	*P* value
Case 1	0.74 ± 0.14	0.99 ± 0.01	0.70	.0027	0.91 ± 0.07	0.97 ± 0.01	0.82	.0039
Case 2	0.84 ± 0.10	0.97 ± 0.01	0.75	.0059	0.74 ± 0.18	0.96 ± 0.04	0.59	.0078
Case 3	0.64 ± 0.23	0.98 ± 0.02	0.54	.0020	0.66 ± 0.22	0.98 ± 0.04	0.55	.0059
Case 4	0.67 ± 0.26	0.99 ± 0.004	0.61	.0029	0.81 ± 0.12	0.98 ± 0.02	0.71	.002
Case 5	0.75 ± 0.30	0.98 ± 0.03	0.63	.0068	0.79 ± 0.19	0.96 ± 0.06	0.64	.0049
Case 6	0.86 ± 0.07	0.98 ± 0.01	0.79	.002	0.78 ± 0.21	0.97 ± 0.03	0.64	.0039
Case 7	0.71 ± 0.11	0.98 ± 0.01	0.66	.00084	0.67 ± 0.27	0.96 ± 0.04	0.54	.0073
Case 8	0.83 ± 0.14	0.95 ± 0.07	0.66	.0038	0.76 ± 0.24	0.95 ± 0.05	0.60	.0029
Case 9	0.79 ± 0.18	0.95 ± 0.05	0.63	.002	0.89 ± 0.08	0.92 ± 0.07	0.67	.0029
Case 10	0.78 ± 0.31	0.94 ± 0.05	0.57	.0017	0.88 ± 0.10	0.94 ± 0.07	0.67	.0071
Mean	0.76 ± 0.18	0.97 ± 0.03	0.65		0.79 ± 0.17	0.96 ± 0.04	0.64	

Abbreviations: CTV = clinical target volume; GTV = gross tumor volume.
Guidelines for kappa measurement analysis (25): <0, poor agreement; 0.01-0.20, slight agreement; 0.21-0.40, fair agreement; 0.41-0.60, moderate agreement; 0.61-0.80, substantial agreement; 0.81-1.00, almost perfect agreement.

volume agreement, sensitivity, and specificity. Analysis of the GTV consensus contours shows the intent to include the complete extent of the gross metastatic tumor using all available clinical information and imaging modalities, including MRI, CT, myelography, plain films, and functional imaging studies such as positron emission tomography CT. Each task force member also

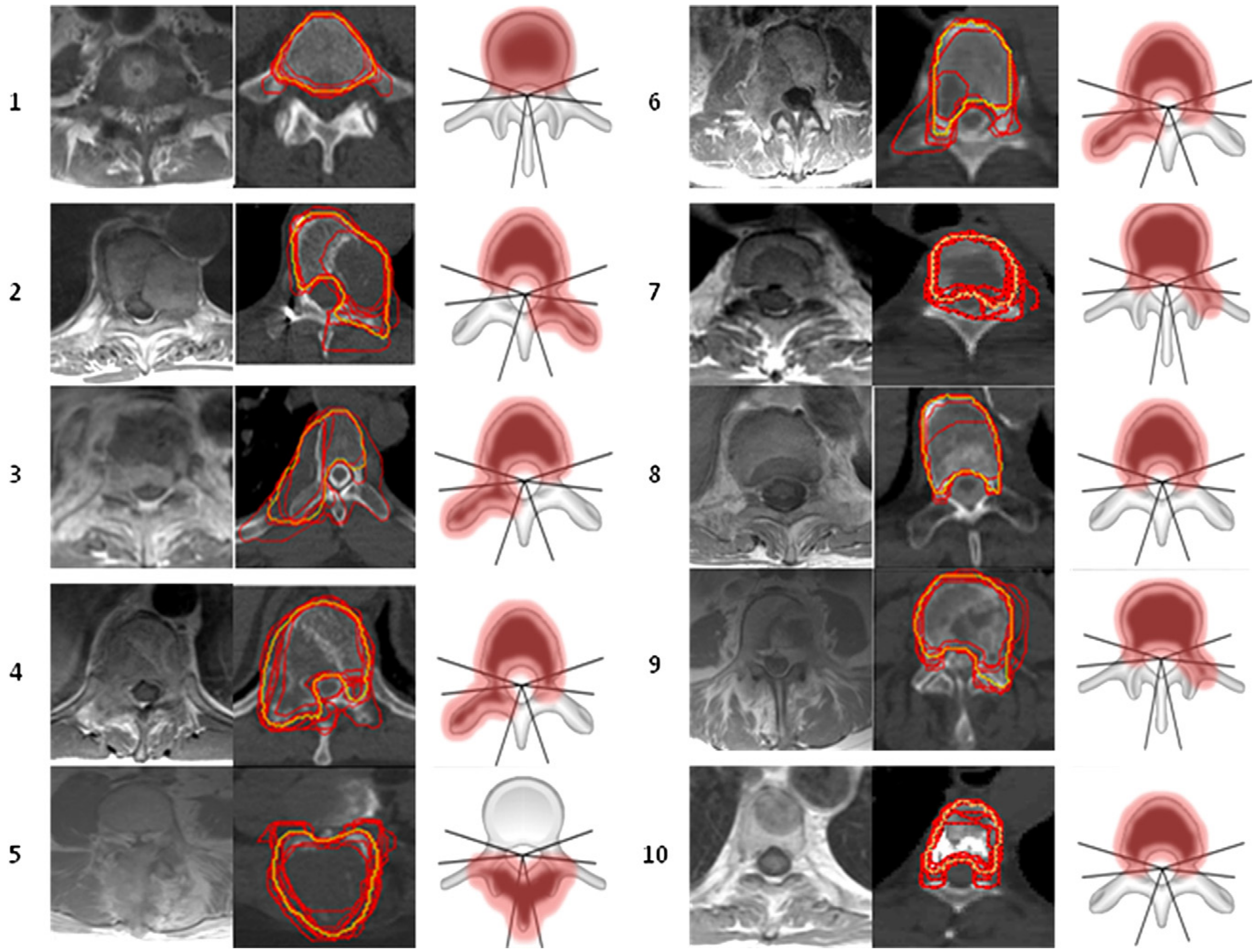


Fig. 2. Consensus clinical target volume contours for spinal stereotactic radiosurgery. Red indicates individual contours and orange indicates consensus contours.

Table 3 Guidelines for spinal SRS bony CTV delineation

GTV involvement	ISRC GTV anatomic classification	ISRC bony CTV recommendation	CTV description
Any portion of the vertebral body	1	1	Include the entire vertebral body
Lateralized within the vertebral body	1	1, 2	Include the entire vertebral body and the ipsilateral pedicle/transverse process
Diffusely involves the vertebral body	1	1, 2, 6	Include the entire vertebral body and the bilateral pedicles/transverse processes
GTV involves vertebral body and unilateral pedicle	1, 2	1, 2, 3	Include entire vertebral body, pedicle, ipsilateral transverse process, and ipsilateral lamina
GTV involves vertebral body and bilateral pedicles/transverse processes	3	2, 3, 4	Include entire vertebral body, bilateral pedicles/transverse processes, and bilateral laminae
GTV involves unilateral pedicle	2	2, 3 ± 1	Include pedicle, ipsilateral transverse process, and ipsilateral lamina, ± vertebral body
GTV involves unilateral lamina	3	2, 3, 4	Include lamina, ipsilateral pedicle/transverse process, and spinous process
GTV involves spinous process	4	3, 4, 5	Include entire spinous process and bilateral laminae

Abbreviations: CTV = clinical target volume; GTV = gross tumor volume; ISRC = International Spine Radiosurgery Consortium.

recommended that all epidural and paraspinal components of tumor be included as a component of the GTV.

CTV contour generation was more complex. All experts did agree on the need to include a bony CTV margin for each case, including abnormal bone marrow signal suspicious for microscopic invasion adjacent to gross tumor in the CTV. Additionally, panelists endorsed the need for an additional bony CTV expansion to encompass adjacent physiologic appearing bone marrow spaces that may harbor subclinical disease and could potentially serve as a nidus for a local recurrence as reported from prospective data (19). Therefore, all consensus CTV contours included an additional margin within the bony compartment to account for subclinical spread of disease. As a general rule, the entire vertebral body, pedicle, transverse process, lamina, or spinous process was included in the CTV if any portion of these regions contained the GTV. Additionally, the next adjacent normal marrow space was typically included in the bony CTV. Specific details and recommendations for bony CTV expansions based on analysis of the consensus contours are given in Table 3. As an example, if the GTV involved the lateral vertebral body and unilateral pedicle, then the CTV would include the entire vertebral body, entire involved pedicle, ipsilateral transverse process, and ipsilateral lamina. There was agreement that no extraosseous CTV expansion (specifically in the epidural or paraspinal soft-tissue spaces) was necessary beyond the GTV in cases of bone-only disease. There was also agreement that circumferential CTVs encircling the cord should be used only in very carefully selected cases, including rare instances in which the vertebral body, bilateral pedicles/lamina, and spinous process are all involved or when there is near-circumferential involvement of the epidural space with metastatic disease.

PTV delineation

PTV contour recommendations are not proposed given the significant differences in inter- and intrafraction motion-management techniques, treatment platforms, immobilization methods, and prescription dose-fractionation schedules currently used at different institutions. Based on survey results, the taskforce did make general recommendations for PTV delineation, including using a uniform 3-dimensional expansion around the CTV, keeping the CTV to PTV margin 3 mm or less, modifying the PTV at the dural margin and adjacent critical structures to allow spacing at the discretion of the treating physician, never overlapping the PTV with the spinal cord or cauda equina, and encompassing the entire GTV and CTV. PTV margins should be generated on a case-by-case basis considering all pertinent variables to ensure adequate CTV target coverage and sparing of adjacent normal tissues.

Summary guidelines

Overall consensus recommendations for GTV, CTV, and PTV delineation in spinal radiosurgery are detailed in Table 4.

Discussion

Approximately 50% of radiation oncologists use spinal radiosurgery in their practices, and spinal treatments represent the second most common SRS treatment site behind that of lung cancers (9). An increasing number of neurosurgeons are also using SRS (1, 10, 11). As spine radiosurgery is more widely applied, a clear need for consensus contouring guidelines has emerged. Standardization of target volumes will enable refinement of radiosurgical practice by allowing meaningful comparisons of outcomes from different institutions, treatment platforms, particle modalities, and dose-fraction schedules. However, current prospective trials in spinal radiosurgery do not used standardized target volume delineation, and some trials (including Radiation Therapy Oncology Group 0631) do not use International Commission on Radiation Units

Table 4 Summary of contouring guidelines for GTV, CTV, and PTV in spinal stereotactic radiosurgery

Target volume	Guidelines
GTV	• Contour gross tumor using all available imaging • Include epidural and paraspinal components of tumor
CTV	• Include abnormal marrow signal suspicious for microscopic invasion • Include bony CTV expansion to account for subclinical spread • Should contain GTV • Circumferential CTVs encircling the cord should be avoided except in rare instances where the vertebral body, bilateral pedicles/lamina, and spinous process are all involved or when there is extensive metastatic disease along the circumference of the epidural space without spinal cord compression
PTV	• Uniform expansion around CTV • CTV to PTV margin $\leq$3 mm • Modified at dural margin and adjacent critical structures to allow spacing at discretion of the treating physician unless GTV compromised • Never overlaps with cord • Should contain entire GTV and CTV

Abbreviations: CTV = clinical target volume; GTV = gross tumor volume; PTV = planning target volume.

nomenclature, which can potentially create confusion over target volume creation and prevent standardized dosimetric reporting. Standardized nomenclature and dosimetric reporting for spine SRS will further improve resident and continuing medical education as the technique of spinal SRS is more widely applied. Similar consensus recommendations have been widely adopted for a variety of organ sites with great success.

Strengths of this study include the high sensitivity and specificity of the consensus contours, indicating substantial agreement between panel members regarding appropriate regions of inclusion and exclusion for the target volumes. The significant *P* values for the CTV and GTV kappa agreement levels demonstrate that the agreement between task force members is not random. By using expectation maximization STAPLE analysis to identify consensus contour recommendations using automated medical image segmentation performance analysis, we overcame the limitations historically associated with human analysis, which is prone to errors from inter- and intraobserver variability. Furthermore, reporting results based on an existing surgical classification scheme will facilitate analysis and discussion of these consensus recommendations.

However, these recommendations do have limitations. Potential sources of interobserver variation in this study include differential handling of the MRI information by each investigator and the use of diverse treatment-planning systems. These guidelines are not intended to be rigid guidelines and should not replace clinical experience or judgment. It is essential to note that although these recommendations were generated from expert consensus, they are not clinically validated nor based on patient outcomes. Also, this analysis focuses on commonly encountered scenarios when spinal SRS is used as *de novo* treatment for spinal metastases and cannot necessarily be applied to the reirradiation or postoperative clinical settings. There were also limited cases of epidural disease extension and scenarios with epidural extension need to be individualized with respect to the degree of epidural space encompassed within the CTV. As with all modern sophisticated radiation treatment-planning techniques, contouring and treatment planning must be completed on a case-by-case basis, with each plan specifically tailored to the patient's individual clinical situation.

Conclusions

This is the first report defining consensus target volume definitions for spinal radiosurgery. These recommendations should serve as a foundation for future refinements in radiosurgery target volume delineation and underscore the need for consensus target definitions in future spinal radiosurgery protocols.

References

1. Bilsky MH, Laufer I, Burch S. Shifting paradigms in the treatment of metastatic spine disease. *Spine (Phila Pa 1976)* 2009;34:S101-S107.
2. Gerszten PC, Mendel E, Yamada Y. Radiotherapy and radiosurgery for metastatic spine disease: what are the options, indications, and outcomes? *Spine (Phila Pa 1976)* 2009;34:S78-S92.
3. Sahgal A, Bilsky M, Chang EL, et al. Stereotactic body radiotherapy for spinal metastases: current status, with a focus on its application in the postoperative patient. *J Neurosurg Spine* 2011;14:151-166.
4. Lovelock DM, Hua C, Wang P, et al. Accurate setup of paraspinal patients using a noninvasive patient immobilization cradle and portal imaging. *Med Phys* 2005;32:2606-2614.
5. Benzil DL, Saboori M, Mogilner AY, et al. Safety and efficacy of stereotactic radiosurgery for tumors of the spine. *J Neurosurg* 2004; 101(Suppl 3):413-418.
6. Ryu SI, Chang SD, Kim DH, et al. Image-guided hypo-fractionated stereotactic radiosurgery to spinal lesions. *Neurosurgery* 2001;49: 838-846.
7. Lutz S, Berk L, Chang E, et al. Palliative radiotherapy for bone metastases: an ASTRO evidence-based guideline. *Int J Radiat Oncol Biol Phys* 2011;79:965-976.
8. Bilsky MH, Laufer I, Fourney DR, et al. Reliability analysis of the epidural spinal cord compression scale. *J Neurosurg Spine* 2010;13: 324-328.
9. Pan H, Simpson DR, Mell LK, et al. A survey of stereotactic body radiotherapy use in the United States. *Cancer* 2011;117:4566-4572.
10. Bilsky MH LI, Fourney D, Groff M, et al. Spinal radiosurgery: A neurosurgical perspective. *J Radiosurg SBRT.* 2011;1:47-54.
11. Sheehan JP. Resident perceptions of radiosurgical training and the effect of a focused resident training seminar. *J Neurosurg* 2010;113:59-63.
12. Yamada Y, Bilsky MH, Lovelock DM, et al. High-dose, single-fraction image-guided intensity-modulated radiotherapy for metastatic spinal lesions. *Int J Radiat Oncol Biol Phys* 2008;71:484-490.

13. Gerszten PC, Burton SA, Ozhasoglu C, et al. Radiosurgery for spinal metastases: clinical experience in 500 cases from a single institution. *Spine (Phila Pa 1976)* 2007;32:193-199.
14. Gibbs IC, Kamnerdsupaphon P, Ryu MR, et al. Image-guided robotic radiosurgery for spinal metastases. *Radiother Oncol* 2007;82:185-190.
15. Ryu S, Rock J, Jain R, et al. Radiosurgical decompression of metastatic epidural compression. *Cancer* 2010;116:2250-2257.
16. Ryu S, James J, Gerszten P, et al. RTOG 0631 Phase II/III study of image-guided stereotactic radiosurgery for localized spine metastases: phase II results. *Int J Radiat Oncol Biol Phys* 2011;81:S131-S132.
17. Garg AK, Shiu AS, Yang J, et al. Phase I/II Trial of single fraction stereotactic body radiotherapy for previously un-irradiated spinal or paraspinal metastases. *Int J Radiat Oncol Biol Phys* 2011;81:S131.
18. Sahgal A, Larson DA, Chang EL. Stereotactic body radiosurgery for spinal metastases: a critical review. *Int J Radiat Oncol Biol Phys* 2008; 71:652-665.
19. Chang EL, Shiu AS, Mendel E, et al. Phase I/II study of stereotactic body radiotherapy for spinal metastasis and its pattern of failure. *J Neurosurg Spine* 2007;7:151-160.
20. Koyfman SA, Djemil T, Burdick MJ, et al. Marginal recurrence requiring salvage radiotherapy after stereotactic body radiotherapy for spinal metastases. *Int J Radiat Oncol Biol Phys* 2011 Oct 8. [Epub ahead of print].
21. Warfield SK, Zou KH, Wells WM. Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation. *IEEE Trans Med Imaging* 2004; 23:903-921.
22. Allozi R, Li XA, White J, et al. Tools for consensus analysis of experts' contours for radiotherapy structure definitions. *Radiother Oncol* 2010;97:572-578.
23. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther* 2005;85: 257-268.
24. Boriani S, Weinstein JN, Biagini R. Primary bone tumors of the spine. Terminology and surgical staging. *Spine (Phila Pa 1976)* 1997;22: 1036-1044.
25. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-174.