Basic Original Report

Stereotactic radiosurgery for multiple brain metastases: Results of multicenter benchmark planning studies

David J. Eaton PhD [a],*, Jonathan Lee MSc [a], Ian Paddick, MSc [b, c]

[a]National Radiotherapy Trials Quality Assurance Group, Mount Vernon Hospital, Northwood, United Kingdom
[b]Medical Physics Ltd, Reading, United Kingdom
[c]Cromwell Hospital, London, United Kingdom

Abstract

Purpose: Stereotactic radiosurgery is indicated for treatment of multiple brain metastases. Various treatment platforms are available, but most comparisons are limited to single-center studies. As part of a national commissioning program, benchmark planning cases were completed by 21 clinical centers, providing a unique dataset of current practice across a large number of providers and equipment platforms.

Methods and materials: Two brain metastases cases were provided, with images and structures predrawn, involving 3 and 7 lesions. Centers produced plans according to their local practice, which were reviewed centrally using metrics for target coverage, selectivity, gradient fall-off, and normal tissue sparing.

Results: Fifty plans were submitted, using 24 treatment platforms. Eleven plans were revised following feedback, including 2 centers that acquired a new platform; 1 other center accepted a restriction of service. All centers prioritized coverage, with the prescription isodose covering ≥95% of 233 of 235 target volumes. Selectivity was much more variable, especially for smaller lesions, and when combined with poor gradient indices resulted in large volumes of normal tissue being irradiated. Tomotherapy submissions were outliers for either selectivity or gradient index, but other platforms could produce plans with relatively low gradient indices for larger lesion volumes. There was more variation among Varian and Elekta LINAC plans than for Gamma Knife and CyberKnife, and larger differences for smaller targets, both inter- and intratreatment platform. Doses to normal brain and brainstem were highest when margins were applied, but improvements were possible by replanning alone.

Conclusions: Multicenter benchmarking exercises have highlighted some variation in clinical practice and priorities, with a few outliers. Most platforms are able to achieve comparable plans, except for the smallest volumes and when larger planning margins are used. The data will be used to advance standardization and quality improvement of national services and can provide useful guidance for centers worldwide.
© 2017 American Society for Radiation Oncology. Published by Elsevier Inc. All rights reserved.

Introduction

Stereotactic radiosurgery (SRS) is characterized by high doses in 1 fraction, highly conformal dose distributions, and high positional accuracy. SRS is indicated for treatment of multiple brain metastases (mets), compared with whole brain radiation therapy.[1,2] Recent evidence supports the noninferiority of SRS without whole brain radiation therapy for up to 10 lesions,[3] and total volume may be a better indication than total number; therefore, use of SRS in this setting is expected to increase, with more centers treating more lesions and the potential expansion of services beyond specialized units. Various treatment platforms are currently available, and plan comparisons have been made between modalities, mostly in single-center studies,[4-8] although a few multicenter studies exist.[9,10] Several have suggested that comparable plans can be generated with linear accelerator (LINAC)-based techniques compared with dedicated units such as Gamma Knife (GK) and CyberKnife (CK). In some cases, however, it is not clear whether these comparisons are biased or represent current clinical practice.

In clinical trials, variation in treatment quality and lack of protocol compliance can cause significant variation in outcome and may even undermine the conclusions of the study.[11] Variation can be minimized, however, by robust quality assurance (QA) programs that are managed by national or regional bodies.[12] Alongside dosimetry audit, benchmark cases are commonly used to assess participating centers using standard cases that are contoured or planned by the center and then analyzed centrally and compared with other submissions.

A prerequisite for all providers selected as SRS centers by the National Health Service in England was to participate in a QA process, informed through collaboration between the national trials QA group (RTTQA) and a multidisciplinary expert advisory group. All clinical centers undertook planning benchmark cases, providing a unique dataset of current practice across a large number of providers and a wide range of equipment. Rather than assessing compliance to a set protocol, no specific guidance was provided, and centers were asked to follow their local practice. This was then used to assess whether the technical competency of each successful bidder was adequate, to facilitate sharing of best practice, to identify outliers, and to support centers with less experience.

Methods and materials

Each provider was required to complete planning benchmark cases, producing treatment plans according to their own local practice, and provide parameters including prescription dose, target conformity, and doses to organs at risk. Two brain mets planning cases were distributed to centers in Digital Imaging and Communications in Medicine radiation therapy format, including computed tomography images, and predrawn structure sets (eFig 1; available as supplementary material online only at www.practicalradonc.org), along with a brief clinical history.

- Case 1: A 67-year-old man with lung cancer completed first-line therapy 10 months previously. He had newly diagnosed extra- and intracranial disease, suitable for second-line systemic therapy, and no previous central nervous system–directed radiation therapy. Posterior fossa metastasis was completely resected, with 3 remaining metastases (including 1 in the brainstem) intended for SRS; volumes provided included gross target volumes (GTVs) (0.1, 0.4, and 0.6 mL), brainstem, optic structures, and brain.
- Case 2: A 65-year-old man who underwent radical resection for lung cancer 10 months previously. Brain magnetic resonance imaging performed after acute presentation with partial seizures revealed 7 lesions, consistent with metastases. He had no other medical history and was otherwise fit. Computed tomography scans showed no extracranial relapse; volumes provided included GTVs (0.1, 0.1, 0.2, 0.3, 0.6, 1.4, and 7.3 mL), brainstem, optic structures, brain, and pituitary fossa.

Centers were asked not to modify any structures, but to add a planning target volume (PTV) margin for positional uncertainty if this was local practice. Centers produced plans for some or all of the lesions, depending on local clinical practice. Doses, volumes, and dose-volume parameters were reported as calculated on their local treatment planning system and used as the primary dataset for this analysis. Structures, plan, and dose cube data were also returned in Digital Imaging and Communications in Medicine radiation therapy format. Independent analysis software VODCA v5.41 (Medical Software Solutions GmbH, Hagendorn, Switzerland) was used to spot-check the submitted values and fill in any gaps or errors.

Several plan quality metrics were calculated following international standard terminology.[13] The first 3 assess the under- or overcoverage of the target by the prescription isodose, with larger values corresponding to better conformity, up to a maximum of 1 in each case. The final 2 metrics assess the surrounding medium dose fall-off. In SRS, steep dose gradients allow high therapeutic doses to be delivered to the target while avoiding or limiting radiation dose to surrounding normal tissue; therefore, a steeper gradient of dose fall-off is normally better, corresponding to lower values of these metrics.

$$\text{Target coverage ratio} = \frac{PTV \ V100\% \ (\text{mL, TTV})}{PTV \ (\text{mL})}$$

$$\text{Selectivity index} = \frac{PTV \ V100\% \ (\text{mL, TTV})}{\text{Total} \ V100\% \ (\text{mL, PIV})}$$

Paddick conformity index (PCI)[14] $= \text{Coverage} \times \text{selectivity}$

Gradient index (GI)[15] $= \dfrac{\text{Total } V50\% \ (\text{mL})}{\text{Total } V100\% \ (\text{mL, PIV})}$

Spread of half prescription isodose,
$\text{R}50\%^{16} = \dfrac{\text{Total } V50\% \ (\text{mL})}{\text{PTV } (\text{mL})}$

V50% or V100% is the volume of a region receiving 50% or 100% of the prescription dose, respectively. Total V100% is commonly known as the prescription isodose volume (PIV) and PTV V100% is also known as treated target volume (TTV). Where no PTV margin was added, GTV volumes were used directly. To assess total tissue doses, regions of interest were created around each lesion so parameters could be calculated individually rather than for the whole patient, which would result in bias toward the larger lesions. Where lesions were inside an OAR, difference volumes were created to assess dose (eg, [brainstem – GTV]).

Results

In total, 50 plans were submitted, using 24 different treatment platforms, as shown in Table 1. Initially, 39 plans were submitted because in some cases local practice was not to use certain platforms for all indications; 11 plans were revised after feedback. Reasons for revision included being outliers in terms of low selectivity, high GI, 2 mm PTV margin and/or high surrounding dose for the brainstem met (case 1), and high normal brain dose (case 2). Both centers

using tomotherapy decided not to use this platform for future treatments but acquired a new platform instead (TrueBeam STx). One other center accepted a restriction of service to 3 mets or fewer, with other patients being referred elsewhere.

Prescription doses were typically 18 to 25 Gy in 1 fraction (or 27 Gy in 3 fractions, 2 plans for case 2), except for the lesion within the brainstem, which was prescribed 12 to 20 Gy in 1 fraction (or 18-30 Gy in 5 fractions, 5 plans for case 1). All centers prioritized coverage, with the prescription isodose covering ≥95% of 233 of 235 target volumes (Fig 1). Selectivity was much more variable, especially for smaller lesions, and in some cases this was combined with a high GI (Fig 2; eFig. 2), resulting in large volumes of normal tissue being irradiated.

Figure 3 shows the relative spread of the half prescription isodose (R50%) against PTV size. Tomotherapy plans gave many of the largest values (range, 5.5-34), whereas GK plans consistently gave the smallest values (range, 2.9-6.0). Other platforms were able to give plans with relatively low R50% for larger lesion volumes, but for smaller volumes there was increasing variation both inter- and intra-treatment platform. There was more variation among Varian and Elekta LINAC plans (ranges, 3.2-13.4 and 3.7-14.4, respectively) than for CK (range, 3.4-8.6).

Doses to normal brain and brainstem also showed some variation by treatment platform, but the greatest impact was the size of PTV margin applied (not seen for GK and CK plans, which all used no margin) (Figs 4-6). Plans initially submitted using 2 mm PTV margin were all outliers and were either revised or, in 1 case, service was restricted. Of the 7 revisions

Table 1 Equipment used for benchmark case submissions, with numbers of platforms shown.

Platform		TPS (version)		Algorithm name	Technique	Collimation	PTV margin (brain lesions)
Gamma Knife	7	Gammaplan (10.1, 11.0)	7	TMR10	Multiple noncoplanar beams	Cones	0 mm
CyberKnife	3	Multiplan (5.21)	3	Ray tracing	Multiple noncoplanar beams	Cones	0 mm (2), 1 mm (2) [a]
Varian LINAC (Novalis / STx / 2100)	8	Eclipse (11.0, 13.6)	2	AAA (1), Acuros XB (1)	Noncoplanar VMAT [b]	2.5 mm MLC	1 mm (2) [a]
		iPlan (4.51-4.54)	5	Pencil beam	Static conformal (1), DCA (3), fixed cone arcs (1)	2.5 mm MLC (4), cones (1)	0 mm (2), 0-1 mm (1), 1 mm (2)
		Pinnacle (9.8)	1	Collapsed cone	Static conformal	2.5 mm MLC	1 mm
Elekta LINAC (Synergy /Agility)	4	Monaco (5.2)	1	Monte Carlo	Noncoplanar VMAT (1)	5 mm MLC	2 mm [a]
		Pinnacle (9.6, 9.8, 14.0)	3	Collapsed cone	Static conformal (1), DCA (1), noncoplanar VMAT (1)	4 mm MLC (2), [c] 5 mm MLC (1)	0 mm (1), 2 mm (2) [a]
Tomotherapy	2	Tomotherapy (5.0, 5.1)	2	Convolution/ superposition	Helical tomotherapy	6 mm binary MLC	0 mm (1), 2 mm (1)

DCA, dynamic conformal arcs; MLC, multileaf collimator; PTV, planning target volume; TPS, treatment planning system; VMAT, volumetric modulated arc therapy.

[a] In 1 of these submissions, the margin was reduced for the brainstem lesion, typically by 1 mm.

[b] 1 Eclipse center used different isocenters for each lesion (case 1 only); all other VMAT submissions used a single isocenter for the whole case.

[c] 1 center used 4 mm MLC for initial submissions and 5 mm MLC for the revision (case 1 only).
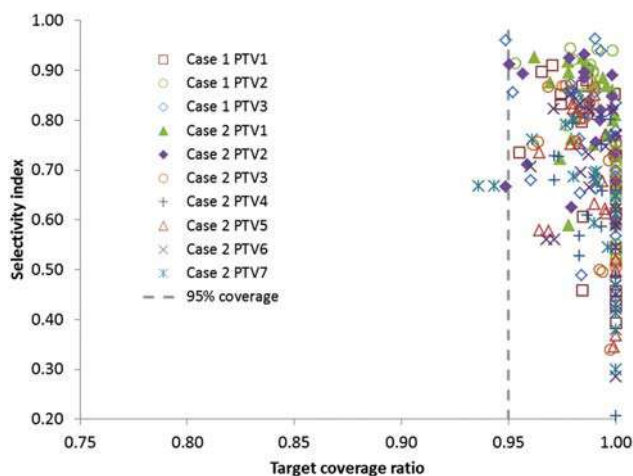
Figure 1 Coverage against selectivity for all lesions, with 95% coverage shown by dotted line. Optimal values are in the top right of this graph. PTV, planning target volume.

using the same platform, 2 reduced the margin on the brainstem lesion (case 1 PTV3, Fig 4), and 1 center increased the inhomogeneity within the targets to increase the gradient fall-off (ie, reduced the prescription isodose). However, 4 plans (2 Varian LINAC centers) made improvements by reoptimizing the plans alone, without changing margin size, technique, or number of fields, based on the feedback on what was achievable for these cases in other centers.

Tomotherapy submissions were outliers in terms of either very high GI (Fig 2) or normal tissue doses (Fig 4; brainstem V10 Gy 6.2 mL; Fig 5, normal brain V12 Gy 148 mL). These centers moved to a new platform as described previously, which accounted for 4 of the 11 revisions.

Discussion

A wide variety of systems and techniques were used, and a wide range of planning parameters was observed; however, prescription doses were similar with almost all

centers using ≥18 Gy for the brain mets, as recommended in a recent review by Lippitz et al.[2]

For brainstem mets, a recent review by Trifiletti et al[17] reported outcome data for 596 brainstem mets from 10 GK centers. Increased toxicity was significantly associated with lesion size, inclusion of whole brain radiation therapy, and marginal (ie, prescription) dose >16 Gy. Prescription doses <16 Gy were associated with reduced local control, but not significantly. They noted that doses of 12 to 20 Gy have been recommended in other studies, which exactly matches the clinical practice found in this study when using no PTV margin. Centers using 1 to 2 mm PTV margins typically prescribed 25 to 30 Gy in 5 fractions to mitigate the impact of the larger volumes.

Four plans used smaller margins for the brainstem lesion than for the other noneloquent mets. This reduces the dose to the normal brainstem by allowing (although not reporting) a degree of undercoverage of this lesion. More generally, Table 1 shows that even different centers with the same equipment applied different margins. GK, CK,
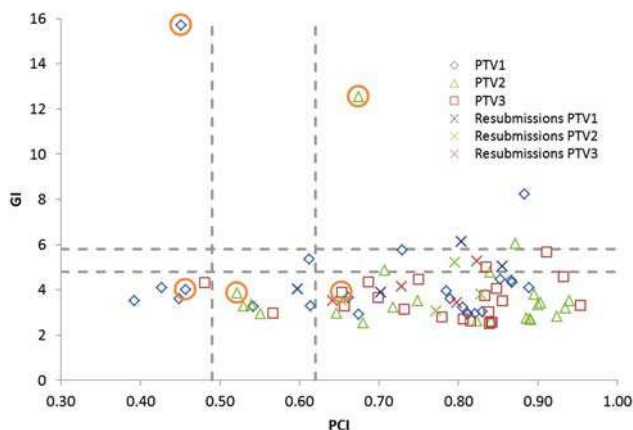
Figure 2 GI against PCI for case 1 (3 mets). Tomotherapy submissions are marked with orange circles, and dotted lines represent 1 and 2 standard deviations from the mean (3.8 and 0.76, respectively), excluding the 2 submissions with GI >10. Optimal values are in the bottom right of the graph. GI, Gradient index; mets, metastases; PCI, Paddick conformity index; PTV, planning target volume.
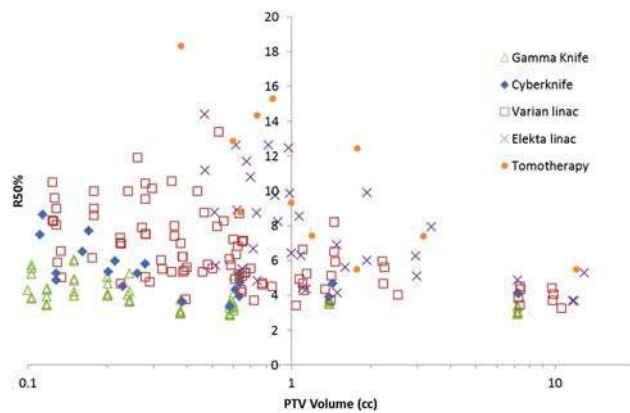
Figure 3   R50% against PTV (logarithmic scale) for all lesions, showing the distribution for each platform. One tomotherapy lesion is not shown off-scale (0.1, 34.0). Optimal values are at the bottom of the graph, corresponding to small values of R50%. PTV, planning target volume; R50%, relative spread of half prescription isodose.

and specialist LINACs such as Varian Novalis or STx all have submillimeter positional accuracy, but some centers took a conservative approach of using a 1 mm PTV margin. Historically, GK centers (and some others in this study) have taken the opposite approach, using no margin and thereby allowing some undercoverage to spare normal tissue. Clinical outcome data include the impact of this philosophy; therefore, even for malignancies, small margins may not be required, unless equipment cannot achieve submillimeter accuracy.

Target conformity

All centers prioritized coverage, presumably to ensure local control; therefore, the PCI index is mainly representative of selectivity (overcoverage). Figure 2 (and eFig 2) shows there is sometimes a tradeoff required between high PCI (minimal overcoverage) or low GI (steep dose falloff); different centers had different priorities in this regard.

For very small lesions, lower selectivity (PCI) is often unavoidable; however, plans in the top left quadrant (low PCI and high GI) were highly undesirable, leading to large volumes of normal tissue being irradiated.

The limited number of tomotherapy plans were overall outliers, particularly in terms of GI or normal tissue doses as described previously, reflecting the spread of low doses probably inherent with a continuous helical delivery and relatively large multileaf collimator (MLC) width (6.25 mm). Soisson et al[18] reported similar conformity and dose fall-off between tomotherapy and circular collimators, so it may be that plans in this study could be further optimized; however, the single center in our study using circular collimators alone (fixed cone arcs) was also an outlier with lower selectivity, so this may be an outdated comparison.

Other studies of brain mets planning have typically compared 2 or 3 modalities, often in a single center, so comparisons are limited. One study reported superior results using noncoplanar volumetric modulated arc therapy (VMAT) compared with GK, but it used an
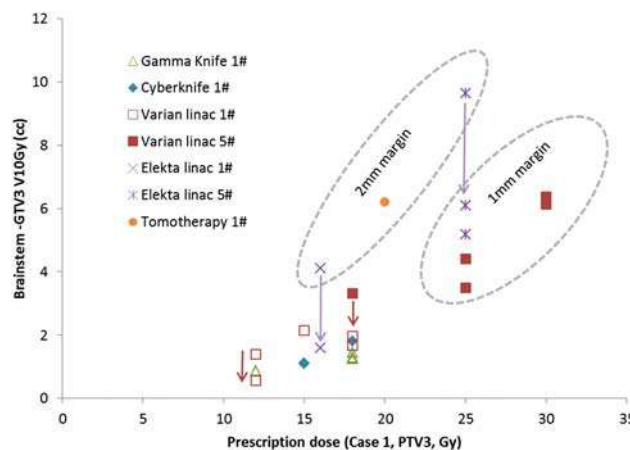
Figure 4   Dose to normal brainstem (brainstem – GTV3 V10Gy) against prescription dose to the brainstem lesion (PTV3) in case 1 (3 mets), for the different fractionations and platforms used. Arrows show the improvement with revision, and dotted ellipses show the plans using 1 and 2 mm PTV margin. All Gamma Knife and CyberKnife plans used no margin. GTV, gross tumor volume; PTV, planning target volume.
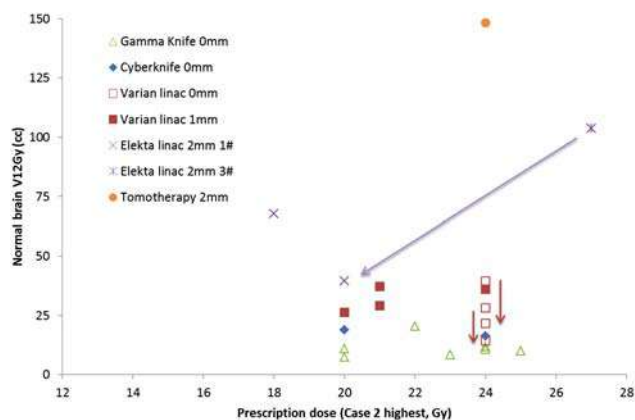
Figure 5 Dose to normal brain (V12Gy) against indicative highest prescription dose (usually given to the smaller lesions) for case 2 (7 metastases), showing the distribution by platform and planning target volume margin used. Arrows show improvement with revision.

older model of the latter.[5] Other VMAT comparisons have typically shown PCI values similar to GK, but higher GI to some extent.[6-8,10] For example, Ma et al[10] compared GK, CK, Novalis dynamic conformal arcs (DCA) (3 mm MLC), and TrueBeam coplanar VMAT (2.5 mm MLC, 2 centers) across 6 centers for 3 to 12 mets in 1 patient. Mean PCI was higher for GK and VMAT (0.74, 0.71) compared with CK and DCA (0.60, 0.49), but normal tissue doses were much higher for VMAT, CK, and DCA. Recently, Gevaert et al[19] compared a new technique for single-isocenter DCA, with (multiple isocenter) DCA and Eclipse VMAT for 10 patients (1-8 mets). PCI was similar among the 3 techniques (mean, 0.65-0.67), but GI was reduced with single isocenter DCA (mean, 3.9 vs 4.5 DCA and 7.1 VMAT).

This study is the largest known planning evaluation of different SRS centers and platforms. When broken down by treatment platform, several trends are apparent (Fig 3). R50% combines the effects of both selectivity and gradient index to give a measure of medium dose conformity, although apparently large increases in this parameter together with

smaller PTV size can actually lead to similar absolute volumes of normal tissue (eg, the same volume of tissue receives at least half of the prescription dose for either R50% 5 and PTV 0.5 mL, or R50% 10 and PTV 0.25 mL).

R50% appears to be correlated with MLC size, with Elekta LINACs (4-5 mm) typically higher than Varian LINACs (2.5 mm). Other studies have suggested larger MLC width contributes to increased spillage and reduction in normal tissue sparing, especially for small brain lesions.[4,6] Our data, however, may be affected by other differences such as PTV margin and planning system (Table 1). Elekta LINAC centers typically applied larger PTV margins to reflect increased uncertainties in isocenter variation with gantry rotation and lack of intrafraction imaging. There are no data points <0.5 mL, but for larger PTVs in case 1 (3 mets), some plans were comparable to other platforms. Elekta centers also used 2 mm dose grid resolution, whereas other centers typically used 1 mm or less, which was the recommendation subsequently given to all centers.
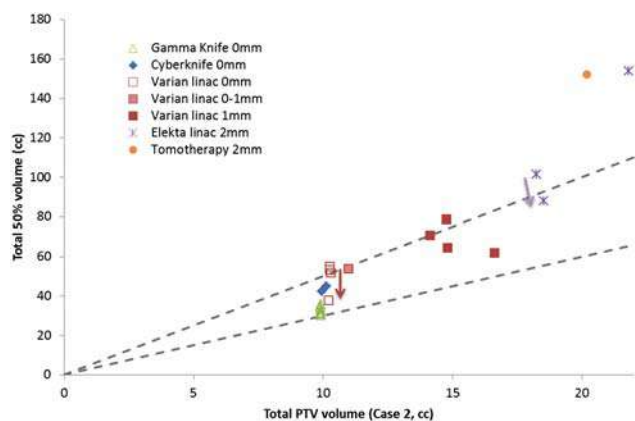
Figure 6 Total V50% against total PTV volume, for case 2 (7 metastases), showing the distribution by platform and PTV margin used. Arrows show improvement for 2 of the 3 revisions, with 1 revision (Varian LINAC 0 mm) not shown because the change was very small. Dashed lines are 3 and 5 times the total PTV volume, which encompassed most of the submissions, with optimal values corresponding to smaller ratios. Variation in PTV for a given margin reflects differences in growth and calculation methods between planning systems. PTV, planning target volume; V50%, volume of 50% of prescription isodose.

The use of VMAT should countermand the limitation of larger MLC size by allowing greater number of degrees of freedom. For case 1, the lowest 2 sets of R50% values for Varian LINAC centers used Eclipse non-coplanar VMAT (albeit with multiple isocenters) and iPlan DCA; for Elekta LINAC centers, the lowest 2 used Monaco noncoplanar VMAT and Pinnacle DCA. For case 2 (7 mets), no Varian VMAT plans were submitted, and the lowest 2 plans used iPlan DCA and Pinnacle static conformal beams. For few mets, therefore, VMAT appears to give comparable results on either platform, but for higher numbers of lesions, further data are required to show the relative merit or detriment. This also applies to a comparison of single versus multiple isocenters, which might be expected to affect the dose fall-off, because all bar one VMAT plan used a single isocenter, but all other LINAC-based plans used multiple isocenters (1 per lesion).

Apparent differences between platforms may also reflect different planning philosophies, such as the level of inhomogeneity within the target volume. Historically SRS plans have "prescribed to the XX% isodose," meaning that the prescription dose is XX% of the maximum dose in the plan (ie, the inverse ratio of the maximum dose to the prescription dose, which is usually quoted in radiation therapy plans). This information was not directly collected; however, a recent national survey found typical values of 45% to 55% for GK, 55% to 80% for CK, and 80% to 100% for other LINAC plans.[20]

These data could be used to derive guidelines for planning metric values; however, they relate to complex multiple mets cases. Only 11 of 235 lesions met suggested optimal values for GK with single targets of PCI >0.85 and GI <3.0.[13,15] In other multiple mets SRS planning studies, the better modality has given PCI values of 0.51 to 0.53,[6] 0.71 to 0.74,[10] and 0.65 to 0.67[19] and GI values of 2.9 to 3.3,[6] 3.7 ± 1.0,[8] and 3.9 ± 1.4.[19] Taking the mean and standard deviation of submissions in this study gives PCI and GI values of 0.62 and 4.8 (3 mets, case 1) and 0.54 and 5.3 (7 mets, case 2). These values could be used as a guide for other centers when planning similar cases, although it should be remembered that low selectivity may be permissible for very small lesions, and normal tissue doses are often a better predictor of toxicity. When applied to simpler cases, optimal values should be achievable, regardless of platform. Individual metrics should always be considered in context when determining the acceptability of a specific treatment plan. Several factors contribute to the quality and accuracy of clinical treatment delivery, and the relative importance of these will vary with equipment and anatomical site.

Normal tissue doses

In spite of differences between platforms in conformity metrics, the greatest impact on normal tissue doses was made by the PTV margin applied. For a lesion within the brainstem, the near-point maximum dose to normal tissue will be very similar to the prescription dose covering the target, so it is more useful to consider the volume irradiated to 10 Gy or more (V10Gy), as shown in Fig 4. Centers with no PTV margin had values of 0.6 to 2.1 mL, with 1 outlier (3.3 mL), which was reduced by replanning. Centers adding a 1 mm margin had higher V10Gy values (3.5-6.4 mL), but this was proportionate to the increased prescription dose (and all used 5 fractions), so can be considered to have equivalent sparing of normal tissue. The 3 centers using a 2 mm margin had disproportionately higher V10Gy values (4.1-9.6 mL), but 2 of these were reduced by using smaller margins (and replanning); the third was a tomotherapy center that subsequently changed platforms.

For normal brain tissue, Flickinger et al[21] were the first to find that V12Gy predicted radionecrosis in brain and that symptoms depended on location. This has since been confirmed by many others, including dependence on target volume and location.[13] For a case simulated with 3 to 12 lesions, Ma et al[10] found that normal brain V12Gy was lowest for GK, whereas for CK and LINAC (DCA or VMAT) plans were 2 to 3 times higher. Gevaert et al[19] reported that normal brain V12Gy for 10 patients with 1 to 8 mets was similar for single or multiple isocenter DCA (36 ± 27 mL) but higher for VMAT (46 ± 36 mL); however, in our study, the greatest impact on normal brain dose is again made by the PTV margin applied (Fig 5).

In a study by Ma et al of 15 clinical targets, replanned on GK with margins between 0.5 and 3 mm, a 2 mm margin increased the target volume by an average of 55% and the risk of symptomatic radionecrosis by 6% to 25%.[22] The detrimental effects of large margins are supported by further clinical studies,[23,24] although prospective studies are still needed. Case 2 submissions using a 1 mm margin had higher normal brain V12Gy values (26-37 mL, compared with 8-28 mL for no margin, following replanning of 2 submissions). The centers using a 2 mm margin had much higher values, more than 10 times the lowest values in one case. One center was able to substantially reduce this volume by more than the reduction in prescription dose, showing the potential to further optimize plans based on what was achievable in other centers. The other 2 centers agreed not to use that platform for more than 3 mets in the future.

These data suggest that the typical single lesion constraint of V12Gy <10 mL should not just be multiplied up by the number of lesions. V12Gy <30 mL has been suggested as a safe level for multiple lesions (M. Yamamoto, private communication). For case 1 (3 mets), the minimum was 1.9 mL, and most centers were <13 mL, except for two 2 mm submissions that were subsequently revised (18 mL, 23 mL). For case 2 (7 mets), the minimum V12Gy was 7.6 mL, and most centers were within 30 mL after replanning. Two 1 mm margin plans and one 2 mm replan were within 40 mL, with the other two 2 mm plans being for platforms that were not to be used in the future. Finally, the total V50% (of prescription)

was typically within 3 to 5 times the total PTV volume, as shown in Fig 6; this metric may form another useful guide for other centers.

Limitations

No guidance was given to centers on how to plan these cases, so rather than assessing compliance, this study is a benchmark of current clinical practice in 1 country. A limited number of cases were used that may not reflect local practice or the typical case mix seen by each provider. It is also not known how much effort each center made to produce optimal plans, although without specific objectives, there is less risk of planners stopping when they have only just met constraints. Conversely, it is possible that centers produced more complex plans than would be practical. For example, a highly conformal plan can be more easily achieved by adding additional isocenters, beams, or arcs. Although these may enhance the planning parameters, they may become unduly burdensome to deliver. For case 2 (7 mets), LINAC-based plans consisted of 45 arcs (fixed cones), 40 to 51 beams (static conformal), 24 to 40 arcs (DCA), or 3 to 5 noncoplanar arcs (VMAT). The improvements seen by reoptimizing plans without changing technique or number of fields demonstrate the benefits of specific guidance on what is achievable for certain cases.

Treatment time is often cited in plan comparison studies as a benefit of single-isocenter VMAT delivery, and it should also be noted that this study considered physical dose only. Protracted treatment times (such as with GK, CK, or many arcs or fields) can lead to reduced biological effectiveness as cells (both tumor and normal tissue) are able to repair sublethal damage. Millar et al[25] reported that, in an extreme case treating vestibular schwannoma with an older model GK, variation of total time from 25 to 130 minutes was associated with reduction of biologically effective dose from 85 $Gy_{2.5}$ to 58 $Gy_{2.5}$ (−32%). For multiple mets, it is the time to treat each lesion that should be considered, which is typically much lower.

Finally, importing and exporting data among such a wide range of systems can lead to errors and corruptions, which were minimized as much as possible. Local treatment planning system values were used where available, unless there were gross discrepancies, in which case VODCA values were used. Uncertainties between calculation of volumes (and hence planning parameters) may lead to variations of 3% to 10%,[26] but these are unlikely to change the overall trends and conclusions in this study.

Conclusion

These benchmarking exercises have highlighted some variation in clinical practice and priorities, with a few outliers that have been removed from clinical use. Most platforms are able to achieve comparable plans, except for smaller volumes and when large margins are used. The data will be used to progress standardization and quality improvement of national services in the future and can provide useful guidance for centers worldwide.

Acknowledgements

References

1. Tsao MN, Rades D, Wirth A, et al. Radiotherapeutic and surgical management for newly diagnosed brain metastasis(es): An American Society for Radiation Oncology evidence-based guideline. *Pract Radiat Oncol.* 2012;2:210-225.

2. Lippitz B, Lindquist C, Paddick I, et al. Stereotactic radiosurgery in the treatment of brain metastases: The current evidence. *Cancer Treat Rev.* 2012;40:48-59.

3. Yamamoto M, Serizawa T, Shuto T, et al. Stereotactic radiosurgery for patients with multiple brain metastases (JLGK0901): A multi-institutional prospective observational study. *Lancet Oncol.* 2014;15:387-395.

4. Wu QJ, Wang Z, Kirkpatrick JP, et al. Impact of collimator leaf width and treatment technique on stereotactic radiosurgery and radiotherapy plans for intra- and extracranial lesions. *Radiat Oncol.* 2009;4:3.

5. Thomas EM, Popple RA, Wu X, et al. Comparison of plan quality and delivery time between volumetric arc therapy (RapidArc) and Gamma Knife radiosurgery for multiple cranial metastases. *Neurosurgery.* 2014;75:409-417.

6. McDonald D, Schuler J, Takacs I, et al. Comparison of radiation dose spillage from the Gamma Knife Perfexion with that from volumetric modulated arc radiosurgery during treatment of multiple brain metastases in a single fraction. *J Neurosurg.* 2014;121(suppl 2):51-59.

7. Huss M, Barsoum P, Dodoo E, et al. Fractionated SRT using VMAT and Gamma Knife for brain metastases and gliomas — a planning study. *J Appl Clin Med Phys.* 2015;16:3-16.

8. Liu H, Andrews DW, Evans JJ, et al. Plan quality and treatment efficiency for radiosurgery to multiple brain metastases: Non-coplanar RapidArc vs. Gamma Knife. *Front Oncol.* 2016;6:26.

9. Ma L, Petti P, Wang B, et al. Apparatus dependence of normal brain tissue dose in stereotactic radiosurgery for multiple brain metastases. *J Neurosurg.* 2011;114:1580-1584.

10. Ma L, Nichol A, Hossain S, et al. Variable dose interplay effects across radiosurgical apparatus in treating multiple brain metastases. *Int J Comput Assist Radiol Surg.* 2014;9:1079-1086.

11. Ohri N, Shen X, Dicker AP, et al. Radiotherapy protocol deviations and clinical outcomes: A meta-analysis of cooperative group clinical trials. *J Natl Cancer Inst.* 2013;106:387-393.

12. Melidis C, Bosch WR, Izewska J, et al. Global harmonization of quality assurance naming conventions in radiation therapy clinical trials. *Int J Radiat Oncol Biol Phys.* 2014;90:1242-1249.

13. Torrens M, Chung C, Chung H-T, et al. Standardization of terminology in stereotactic radiosurgery: Report from the Standardization Committee of the International Leksell Gamma Knife Society. *J Neurosurg.* 2014;121(suppl 2):2-15.

14. Paddick I. A simple scoring ratio to index the conformity of radiosurgical treatment plans. *J Neurosurg.* 2000;93(suppl 3): 219-222.

15. Paddick I, Lippitz B. A simple dose gradient measurement tool to complement the conformity index. *J Neurosurg.* 2006;105:194-201. [Suppl.].

16. Timmerman R, Paulus R, Galvin J, et al. Stereotactic body radiation therapy for inoperable early stage lung cancer. *JAMA.* 2010;303: 1070-1076.

17. Trifiletti DM, Lee C-C, Kano H, et al. Stereotactic radiosurgery for brainstem metastases: An international cooperative study to define response and toxicity. *Int J Radiat Oncol Biol Phys.* 2016;96:280-288.

18. Soisson ET, Mehta M, Tome W. A comparison of helical tomotherapy to circular collimator-based linear-accelerator radiosurgery for the treatment of brain metastases. *Am J Clin Oncol.* 2011;34:388-394.

19. Gevaert T, Steenbeke F, Pellegri L, et al. Evaluation of a dedicated brain metastases treatment planning optimization for radiosurgery: A new treatment paradigm? *Radiat Oncol.* 2016;11:13.

20. Dimitriadis A, Kirkby KJ, Nisbet A, et al. Current status of cranial stereotactic radiosurgery in the UK. *Br J Radiol.* 2016;89:20150452.

21. Flickinger JC, Kondziolka D, Pollock BE, et al. Complications from arteriovenous malformation radiosurgery: Multivariate analysis and risk modeling. *Int J Radiat Oncol Biol Phys.* 1997;38:485-490.

22. Ma L, Sahgal A, Larson DA, et al. Impact of millimeter-level margins on peripheral normal brain sparing for gamma knife radiosurgery. *Int J Radiat Oncol Biol Phys.* 2014;89:206-213.

23. Nataf F, Schlienger M, Liu Z, et al. Radiosurgery with or without A 2-mm margin for 93 single brain metastases. *Int J Radiat Oncol Biol Phys.* 2008;70:766-772.

24. Kirkpatrick JP, Wang Z, Sampson JH, et al. Defining the optimal planning target volume in image-guided stereotactic radiosurgery of brain metastases: Results of a randomized trial. *Int J Radiat Oncol Biol Phys.* 2015;91:100-108.

25. Millar WT, Hopewell JW, Paddick I, et al. The role of the concept of biologically effective dose (BED) in treatment planning in radiosurgery. *Phys Med.* 2015;31:627-633.

26. Ma L, Sahgal A, Nie K, et al. Reliability of contour-based volume calculation for radiosurgery. *J Neurosurg.* 2012;117:203-210. [Suppl.].