

# Improving accuracy for stereotactic body radiotherapy treatments of spinal metastases

James Rijken<sup>1,2</sup> | Barry Jordan<sup>1</sup> | Scott Crowe<sup>2,3</sup> | Tanya Kairn<sup>2,3</sup> | Jamie Trapp<sup>3</sup>

<sup>1</sup>Genesis Care, Flinders Private Hospital, Bedford Park, SA, Australia

<sup>2</sup>Queensland University of Technology, Brisbane, QLD, Australia

<sup>3</sup>Royal Brisbane and Women's Hospital, Brisbane, QLD, Australia

Author to whom Correspondence should be addressed. James Rijken

E-mail: james.rijken@adradcentre.com.au

## Abstract

**Purpose:** Use of SBRT techniques is now a relatively common recourse for spinal metastases due to good local control rates and durable pain control. However, the technique has not yet reached maturity for gantry-based systems, so work is still required in finding planning approaches that produce optimum conformity as well as delivery for the slew of treatment planning systems and treatment machines.

**Methods:** A set of 32 SBRT spine treatment plans based on four vertebral sites, varying in modality and number of control points, were created in Pinnacle. These plans were assessed according to complexity metrics and planning objectives as well as undergoing treatment delivery QA on an Elekta VersaHD through ion chamber measurement, ArcCheck, film-dose map comparison and MLC log-file reconstruction via PerFraction.

**Results:** All methods of QA demonstrated statistically significant agreement with each other ( $r = 0.63$ ,  $P < 0.001$ ). Plan complexity and delivery accuracy were found to be independent of MUs ( $r = 0.22$ ,  $P > 0.05$ ) but improved with the number of control points ( $r = 0.46$ ,  $P < 0.03$ ); with use of 90 control points producing the most complex and least accurate plans. The fraction of small apertures used in treatment had no impact on plan quality or accuracy ( $r = 0.29$ ,  $P > 0.05$ ) but rather more complexly modulated plans showed poorer results due to MLC leaf position inaccuracies. Plans utilizing 180 and 240 control points produced optimal plan coverage with similar complexity metrics to each other. However, plans with 240 control points demonstrated slightly better delivery accuracy, with fewer MLC leaf position discrepancies.

**Conclusion:** In contrast to other studies, MU had no effect on delivery accuracy, with the most impactful parameter at the disposal of the planner being the number of control points utilized.

## PACS

87.55.-x

## KEY WORDS

accuracy, control point, SABR, spine, VMAT

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2018 Genesis Care. *Journal of Applied Clinical Medical Physics* published by Wiley Periodicals, Inc. on behalf of American Association of Physicists in Medicine.

## 1 | INTRODUCTION

Stereotactic body radiotherapy (SBRT) describes extracranial treatment techniques which utilize a larger delivery of radiation dose than conventional radiotherapy and in fewer fractions, resulting in a higher biological effective dose for the treatment site.<sup>1</sup> As with stereotactic radiosurgery for brain metastases,<sup>2</sup> the SBRT technique is able to provide extra dose to the target volume without exceeding recommended normal tissue tolerances. In the case of spinal metastases, the spinal cord may have already been irradiated through conventional radiotherapy and SBRT provides a noninvasive treatment option.<sup>3</sup>

In order to limit dose to the spinal cord, modulated treatment techniques are used to produce complex dose distributions that spare organs at risk.<sup>4–6</sup> SBRT spine treatment can be delivered using a number of treatment machines like helical tomotherapy units, robotic radiosurgery systems or on a linac utilizing intensity-modulated radiotherapy (IMRT) or volumetric-modulated arc therapy (VMAT). In an Eclipse treatment planning system (TPS) & Varian linac environment, single arc VMAT has been found to produce inferior target coverage and normal tissue sparing compared to IMRT, but two arc VMAT is able to produce comparable results.<sup>7</sup> Comparisons between flattening filter free (FFF) and conventional modalities have demonstrated significant improvements in normal tissue/spinal cord sparing for FFF plans with a greater number of control points, while achieving the same level of target coverage.<sup>8,9</sup> While these and other planning studies have been conducted for SBRT spine, there are still gaps in the literature regarding the overall optimization of the SBRT spine technique, which requires further investigation.

A couple of planning studies have also included quality assurance (QA). A recent study found accuracy improvements for an individual SBRT spine case when the gantry spacing between control points was decreased from 4° to 3°. The delivery accuracy of different IMRT and VMAT techniques has also been investigated for spinal treatment.<sup>11</sup> IMRT was measured to deliver more accurately than VMAT due to its reduced complexity. VMAT beams can be delivered more accurately if optimized with a monitor unit (MU) limit, which can produce a less complex plan. However, these findings may be unique to the dose calculation and optimization algorithms of the Eclipse TPS and the accuracy of the Varian linac setup. The study also allowed a slightly higher cord dose which could make large impacts in plan conformity. Another recent study investigated the effect of high definition MLCs on plan quality, complexity and deliverability<sup>12</sup> and found that while slight improvements to quality could be made for VMAT SBRT spine treatments, plan complexity increased and thus delivery accuracy suffered.

There is an overall lack of studies which combine both planning optimizations and QA results for SBRT spine treatment. This study aims to address this and improve upon previous studies through utilization of four different methods of QA. It is important to fully assess inaccuracies/accuracies in the TPS, so that it can be used confidently in the correct range of parameters.

## 2 | MATERIALS AND METHODS

Four previously treated and anonymized SBRT spine CT and contour datasets were selected with differing vertebra and target volume shapes (targeting had been previously performed by radiation oncologists according to international guidelines,<sup>3,13</sup> CT was performed with 2 mm slices). Eight different SBRT spine plans were created on each CT image dataset, utilizing both the 6 MV and 6 MV FFF beam. Control points were set as planning constraints before inverse optimization and plans of 50, 90, 180, and 240 control points were created (Table 1). The 50 control point plans consisted of nine beams delivering IMRT, while the 90, 180, and 240 control point plans were 1, 2, and 2 arc VMAT plans, respectively. Plan complexity was measured through different metrics: (a) modulation index (MI) which considers the fluence map (higher MI implies higher complexity),<sup>14</sup> (b) modulation complexity score (MCS) which considers aperture area variability and leaf sequence variability (higher MCS implies lower complexity),<sup>15</sup> and (c) the small aperture score (SAS) which considers the fraction of fields used <10 mm (higher SAS corresponds to more small fields).<sup>16</sup>

The TPS utilized was Pinnacle3® 9.10 (Koninklijke Philips N.V., Amsterdam, The Netherlands). Plans were created within the TPS each with the same fractionation and dose constraints currently used clinically: 30 Gy in 3 fractions, with 100% of planned target volume (PTV) to be covered by 80% of prescribed dose,<sup>17</sup> but ideally covered by 90% as per RTOG 0631.<sup>13</sup> The other main metric concerned was assessment of the Paddick conformity index (CI) for target coverage (CI <1.0 implies poorer conformity)<sup>18</sup> as per ICRU 91.<sup>19</sup> Other departmental metrics concerned were PTV mean dose below 110% of prescription and 95% of the clinical target volume (CTV) to receive 98% of the prescription<sup>1</sup> (CTV-PTV margin of 2 mm). However, critical nervous structure (CNS) constraints always took priority. The main CNS structure concerned was the thecal sac with a maximum dose of 20 Gy.<sup>20</sup>

Each of the 32 clinical treatment plans created was given the same iteration schedule for inverse planning. The dose calculations were made using the collapsed cone convolution (CCC) algorithm on a 2 mm dose grid<sup>21</sup> with a minimum leaf separation of 0.5 cm. The optimization schedule began with an initial 80 iterations after which the plan was assessed and the optimization objectives adjusted in order to achieve the desired conformity. After these adjustments,

**TABLE 1** List of SABR spine treatment plans

Vertebral region	Planning volume	Energy	Control points
T12	Vertebral body	6 MV	50, 90, 180, 240
		6 MV FFF	50, 90, 180, 240
L2	Whole vertebra	6 MV	50, 90, 180, 240
		6 MV FFF	50, 90, 180, 240
L1	Vertebral body	6 MV	50, 90, 180, 240
		6 MV FFF	50, 90, 180, 240
T11	Whole vertebra	6 MV	50, 90, 180, 240
		6 MV FFF	50, 90, 180, 240

the plan was then optimized for a further 30 iterations. The plan was adjusted and optimized another two times, bringing the total number of iterations to 170 with three adjustments.

The treatment plans were delivered on an Elekta VersaHD<sup>®</sup> linac (Elekta, Stockholm, Sweden) commissioned down to field sizes of 1 × 1 cm<sup>2</sup> with a RAZOR<sup>™</sup> Chamber (IBA Dosimetry GmbH, Schwarzenbruck, Germany) and an EDGE Detector<sup>™</sup> (Sun Nuclear Corp., Melbourne, FL, USA). The first mode of QA was through isocenter ion chamber (IC) measurement in a solid water slab using the calibrated RAZOR<sup>™</sup> Chamber (calibration traceable to primary standards laboratory). The dose measured by the small volume IC was compared to the TPS calculation. Second, the plans were delivered to an ArcCheck<sup>®</sup> device (Sun Nuclear Corp., Melbourne, FL, USA). Third, film QA was performed in the sagittal plane, utilizing the ArcCheck as a phantom. The film employed was Gafchromic<sup>™</sup> EBT3 (Ashland Specialty Ingredients, Bridgewater, NJ, USA) and was scanned on an Expression<sup>™</sup> 10000XL scanner (Epson<sup>®</sup>, North Ryde, NSW, Australia) and processed through SNC Patient Software 6.1 (Sun Nuclear Corp., Melbourne, FL, USA). A film-dose calibration was also performed at the time of delivery such that SNC Patient was able to handle exposures between 1 and 20 Gy. The dose maps derived

from the film (red channel, based on the work of Micke et al.<sup>22</sup>) and the ArcCheck were then compared to respective dose maps from the TPS through gamma analysis in SNC Patient with respective gamma criteria of 3%/1.5 mm and 2%/2 mm. All gamma analyses in this study were performed through absolute dose comparison with a local calculation and a 10% low dose cut-off threshold.<sup>23</sup>

Finally, QA on the treatment plan delivery was also performed through the PerFraction<sup>™</sup> software package (Sun Nuclear Corp., Melbourne, FL, USA). PerFraction takes the MLC log-files from the linac and reconstructs the dose delivered on the CT image dataset using its own CCC style algorithm. A gamma analysis of the two calculated dose distributions was performed (1%/1 mm) for each plan. Correlation of data was assessed using the Pearson correlation coefficient,  $r$ ,<sup>24</sup> given by

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

where  $n$  is the number of samples and  $x_i$  and  $y_i$  are the samples. The statistical significance of any correlation was assessed through  $P$  values calculated using a Student's  $t$  test ( $\alpha = 0.05$ ).

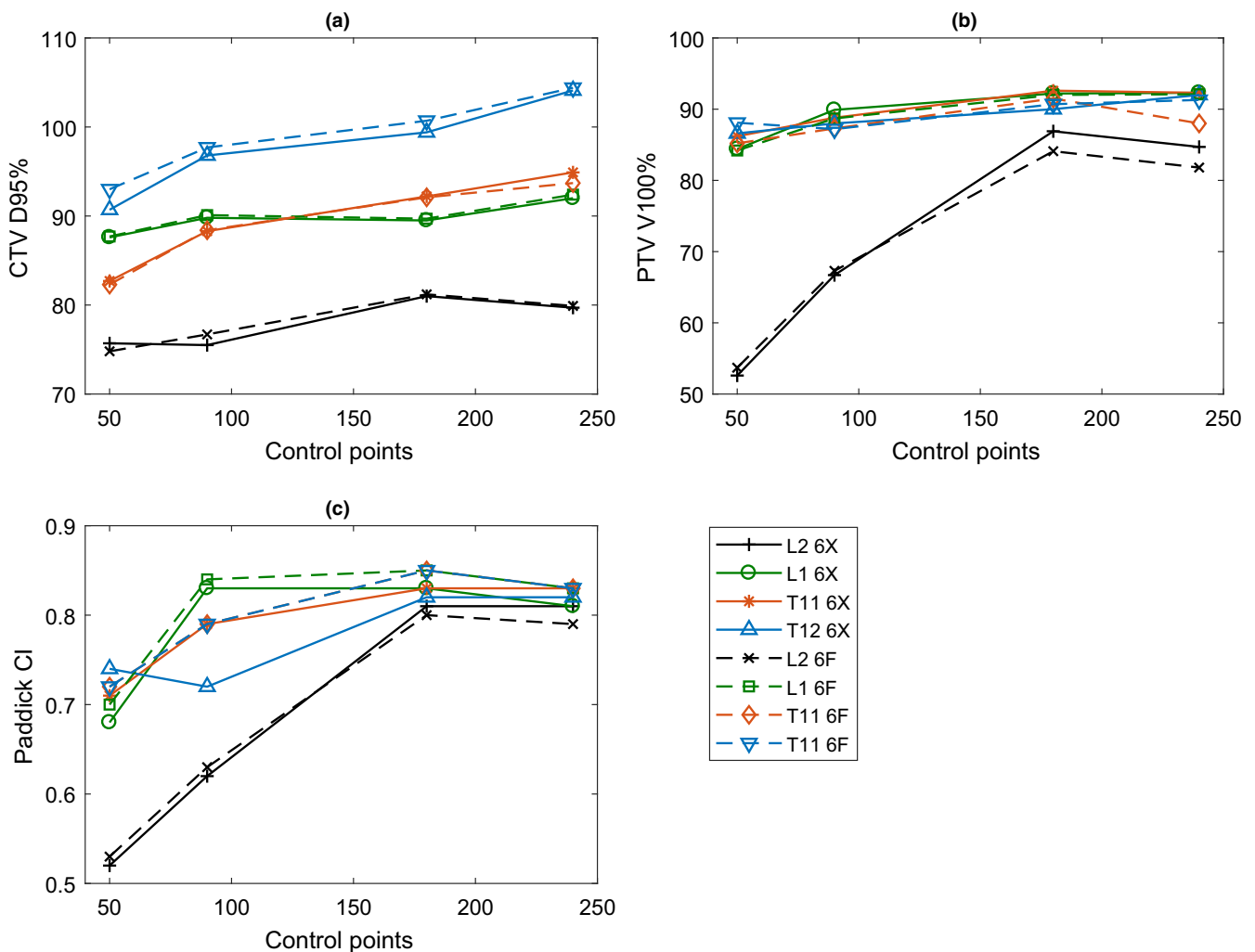


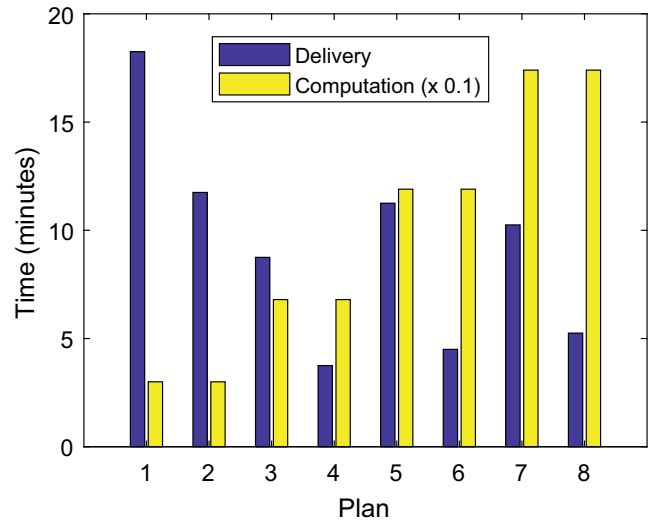
FIG. 1. Plan quality metrics plotted against number of control points: (a) CTV D95%, (b) PTV V100%, and (c) Paddick CI.

### 3 | RESULTS

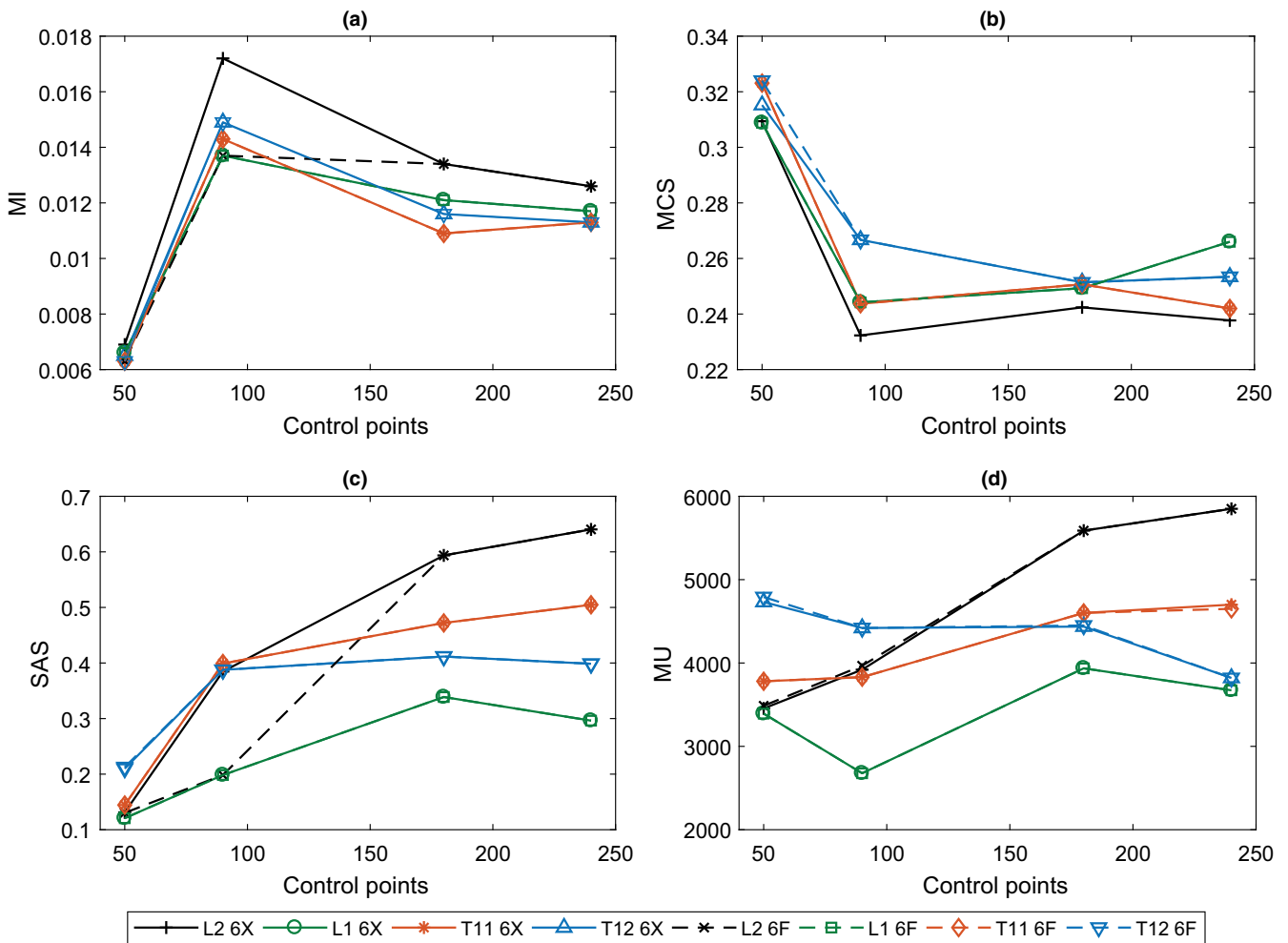
Plans are labeled in the results according to their site (L1, L2, T11, or T12) and/or control points (50, 90, 180, or 240) and/or modality ("X" for 6 MV or "F" for 6 MV FFF). Plan quality results of the SBRT spine plans can be seen in Fig. 1. The most important planning criterion was the CNS dose constraint with the resultant median thecal sac dose being 18.7 Gy. The mean Paddick CI for 6 MV plans was  $0.76 \pm 0.09$  and for 6 MV FFF was  $0.77 \pm 0.09$ . The impact of prioritizing CNS sparing can be seen in the difficulty to achieve good conformity and coverage (median CTV D95% was 89.9%, median PTV V100% was 88.0%, median PTV mean dose was 110.8%) in most cases. However, as the number of control points was increased, the coverage and conformity consistently improved ( $r = 0.70, P < 0.001$ ). There was no significant correlation between plan quality and the number of MUs ( $r = 0.16, P > 0.05$ ).

MCS and MI had no significant trend with MU ( $r = 0.22, P > 0.05$ ). However, Fig. 2 shows both varied with the number of control points, demonstrating the same changes in plan complexity. The 50 control point IMRT plans were the least complex and the 90 control point VMAT were the most, with 180 and 240 being of comparable complexity. SAS increased with the number of control points

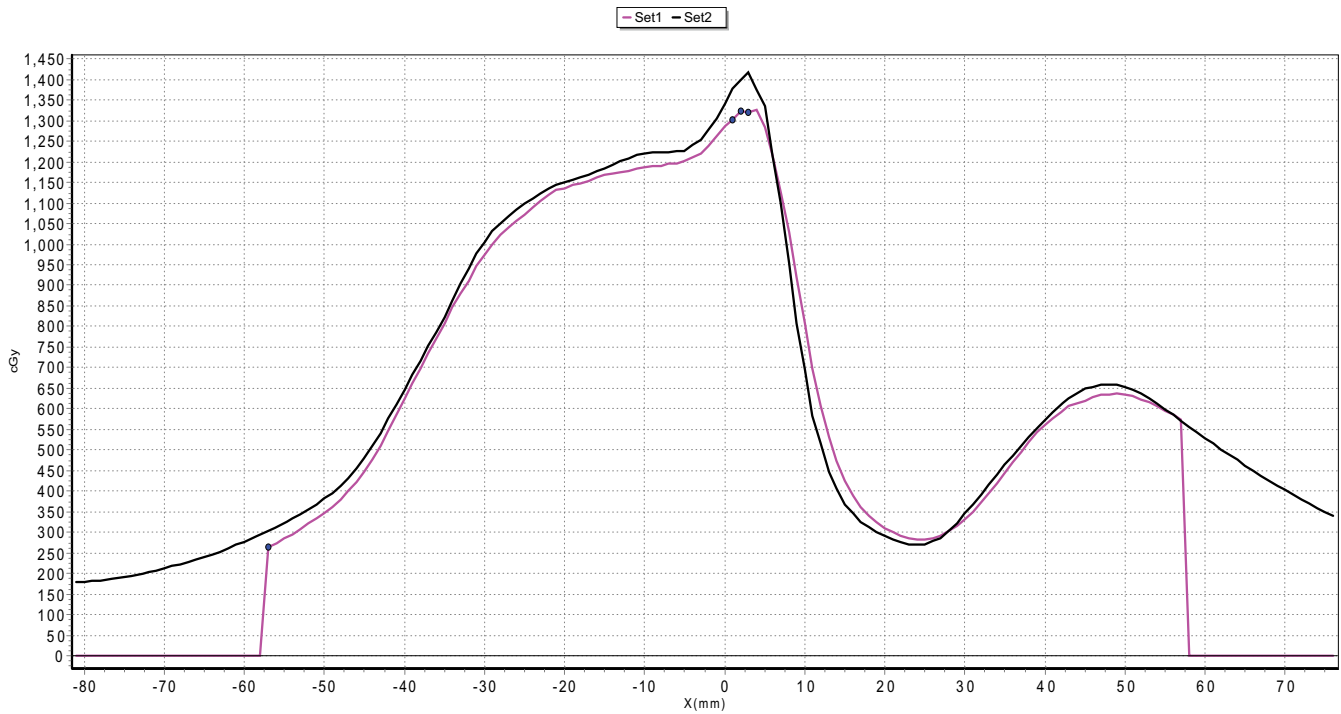
used, with 180 and 240 control point plans having comparable SAS in most plans. No consistent trend was found between MU and the number of control points.



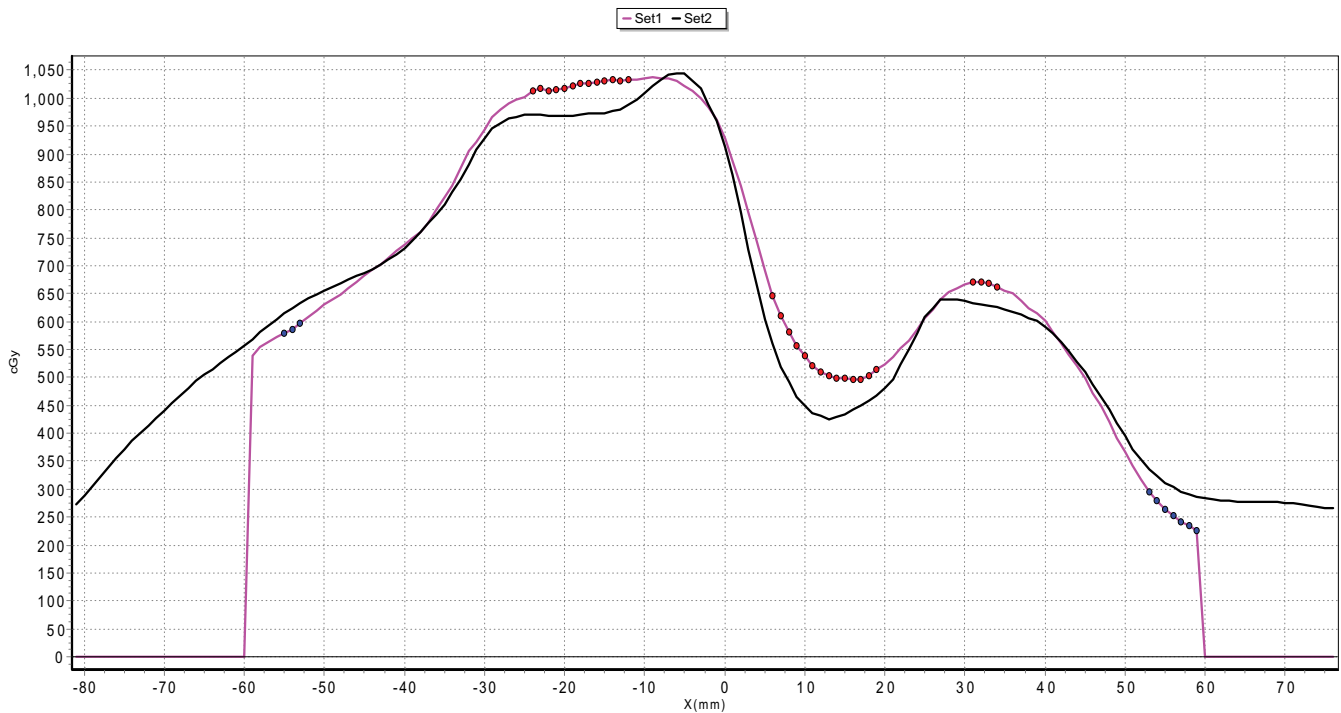
**Fig. 3.** Average times taken for treatment planning and delivery for each plan type: (a) 50X, (b) 50F, (c) 90X, (d) 90F, (e) 180X, (f) 180F, (g) 240X, (h) 240F.



**Fig. 2.** Plan metrics for each plan plotted against the number of control points: (a) MI, (b) MCS, (c) SAS, and (d) MU.



**FIG. 4.** Film QA profile from the T11 240X plan with delivered dose (set 1) and planned dose (set 2). Gamma pass rate of 98.2% (3%/1.5 mm).

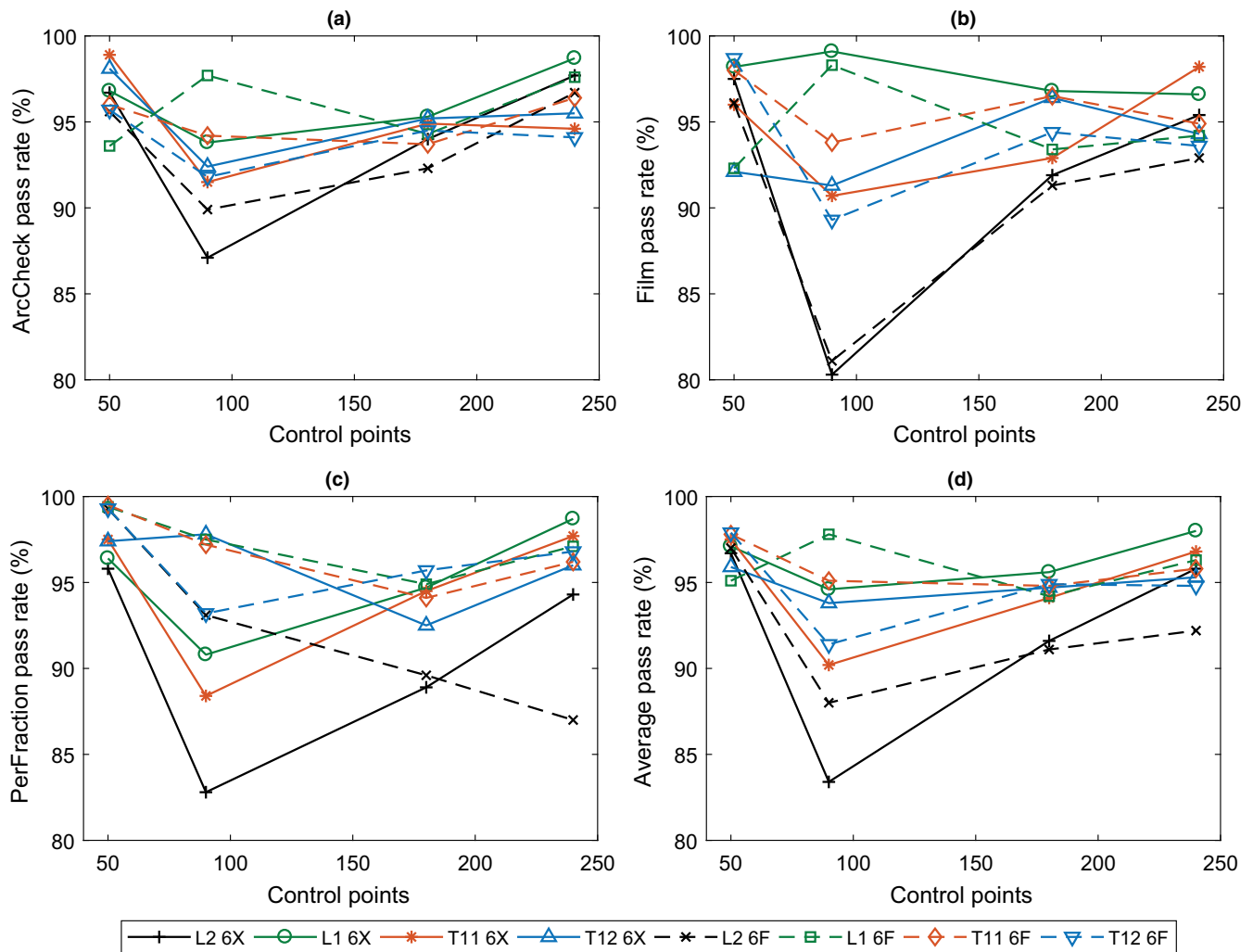


**FIG. 5.** Film QA profile from the L2 90X plan with delivered dose (set 1) and planned dose (set 2). Gamma pass rate of 80.3% (3%/1.5 mm).

Average planning computation times and treatment delivery times are shown in Fig. 3. Computation times scale according to the number of control points utilized. Plan delivery times were quicker for FFF due to the higher dose rate and increased with the number of arcs or beams used.

As an example, a plan that was delivered with a high level of accuracy was T11 240X with its film QA shown in Fig. 4, while a particularly poorly delivered plan, L2 90X, is shown in Fig. 5.

ArcCheck, Film, PerFraction, and mean pass rates plotted against control points are shown in Fig. 6. Film results would indicate that



**FIG. 6.** Gamma pass rates from QA for each plan through (a) ArcCheck (2%/2 mm), (b) film (3%/1.5 mm), (c) PerFraction 3D analysis (1%/1 mm), and (d) mean gamma pass rates plotted against the number of planned control points.

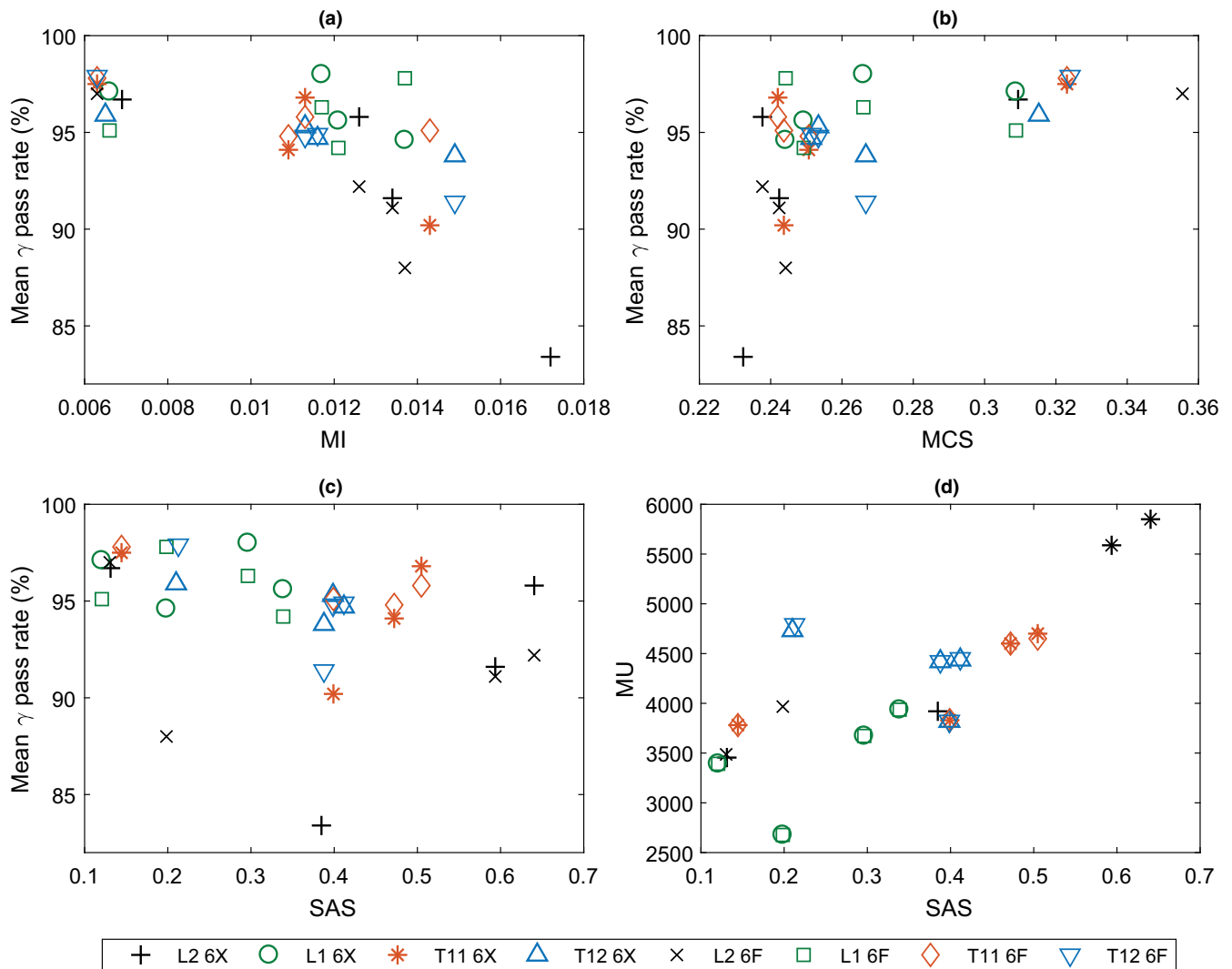
there is no consistent difference between 240 and 180 control point plans but combined with ArcCheck and PerFraction results, a statistically significant improvement in gamma pass rates was observed from 90 to 240 control points ( $r = 0.46$ ;  $P < 0.03$ ). The median respective pass rates for ArcCheck, Film, and PerFraction were 95.1%, 94.4%, and 95.9% at their respective gamma criteria. All methods of gamma analysis demonstrated statistically significant agreement with each other ( $r = 0.63$ ,  $P < 0.001$ ) with no such trend observed between pass rate and MUs ( $r = 0.18$ ,  $P > 0.05$ ). 50 and 240 control point plans were found to be the most accurate, with plans of 90 control points performing the most poorly and with the largest variation in results.

Figure 7 shows the relationship between plan accuracy and plan complexity. (a) and (b) demonstrate that as plan complexity decreases, gamma rates become consistently better ( $r = 0.51$ ,  $P < 0.02$ ). (c) shows the lack of correlation between gamma pass rates and SAS for each plan ( $r = 0.29$ ,  $P > 0.05$ ) and (d) shows the relationship between SAS and MUs, demonstrating a statistically significant trend ( $r = 0.79$ ,  $P < 0.001$ ).

MCS and MI had no significant trend with MU ( $r = 0.22$ ,  $P > 0.05$ ) and while the median IC result was a  $-0.05\%$  difference to TPS, isocenter dose differences became more negative as MU increased ( $r = 0.46$ ,  $P < 0.01$ ). Correlation data have been collated in Table 2.

## 4 | DISCUSSION

Given the CNS prioritization, it is not surprising that very few plans achieve the target goals (Fig. 1). However, such plans are still clinically deliverable as 27 Gy in 3 fractions. Contrary to other studies, no difference was found in conformity between 6 MV and 6 MV FFF plans,<sup>9</sup> with IMRT plans consistently less conformal compared to 90 control point VMATs and poorer still compared to 180 and 240.<sup>7</sup> No correlation could be found between the metrics of plan complexity and quality. These metrics are not under direct control of the user anyway, apart from SAS which can be influenced in the model. The most impactful parameter at the disposal of the planner



**FIG. 7.** Plan delivery accuracy as a measure of mean gamma pass rate plotted against (a) MI, (b) MCS, and (c) SAS. (d) shows the relationship between MU and SAS.

is the number of control points utilized, with major improvements made to plan quality and consistency of quality. Plans with 50 control points composed the worst plans with their poor conformity attributed to the use of IMRT and its limitations.<sup>25,26</sup> No statistically significant trend could be observed between conformity and MUs, which is in disagreement with results shown by Kairn et al. for the Varian/Eclipse setup, where VMAT plans optimized with MU limitations resulted in poorer conformity.<sup>11</sup>

It can be seen that better quality plans only consistently resulted from increases to the number of control points (90 to 240), so it is of interest then that as control points increased, the likelihood of producing a plan with a greater SAS also increased [Fig. 2(c)]. Without fully analyzing the impact of SAS on plan quality, one might be tempted to deduce that small fields produce better quality plans — this was found not to be the case ( $r = 0.15, P > 0.05$ ). The most complex plans resulted from setting control points to 90, causing the VMAT optimization algorithm to produce more complicated MLC shapes in order to meet planning objectives. IMRT plans were not as

complex, even though there were fewer control points, due to optimization algorithm differences and as such, the delivery accuracy was high.

The relative improvements in treatment times are similar to those seen in other studies,<sup>8,27–29</sup> but not so severe as those shown by Kairn et al. where IMRT plans took around three times as long to deliver.<sup>11</sup> This is the first report of average TPS computation times for a SBRT spine study.

The significant agreement between gamma pass rates from three methods of QA (Fig. 6) may be intuitive for comparisons between film and ArcCheck results (if one has refined their QA procedure), but PerFraction considers the MLC log-files when assessing plan delivery, so the agreement is a little more subtle. The correlation, then, between ArcCheck/film and PerFraction is of interest and has not yet been reported for any MLC log-file-based system for SBRT spine. Past investigations of PerFraction have been limited to its EPID dosimetry capabilities<sup>30–32</sup> and some work has been performed on other MLC log-file-based systems, demonstrating good agreement



**TABLE 2** Pearson correlation coefficients and statistical significance between experiment data

Variable 1	Variable 2	r	P	Comment
Plan quality	MU	0.16	>0.05	No significant correlation
Plan quality	Control points	0.70	<0.001	Significant correlation
Plan complexity	MU	0.22	>0.05	No significant correlation
$\gamma$ pass rate	MU	0.18	>0.05	No significant correlation
$\gamma$ pass rate	SAS	0.15	>0.05	No significant correlation
$\gamma$ pass rate	Control points (90–240)	0.46	<0.03	Significant correlation
$\gamma$ pass rate	Plan complexity	0.51	<0.002	Significant negative correlation
MU	Plan complexity	0.51	<0.002	Significant negative correlation
$\gamma$ pass rate	SAS	0.29	>0.05	No significant correlation
MU	SAS	0.79	<0.001	Significant correlation
Isocenter dose	MU	0.46	<0.01	Significant negative correlation
ArcCheck pass rates	Film pass rates	0.75	<0.001	Significant correlation
ArcCheck pass rates	PerFraction pass rates	0.58	<0.001	Significant correlation
Film pass rates	PerFraction pass rates	0.56	<0.001	Significant correlation

between conventional QA and MLC log-file-based patient QA.<sup>33–36</sup> Discrepancies have been demonstrated between actual and intended MLC leaf delivery for low MU control points, with good agreement shown between MLC log-file and diode array assessment methods.<sup>37</sup> A similar effect, in this study, may be attributed to the correlation between PerFraction and ArcCheck/film, thus inferring that losses in plan accuracy are due to MLC movement discrepancies. While most of the PerFraction gamma fail points were in peripheral tissue and neither consistently higher nor lower in dose, target and CNS doses were generally greater than the TPS. This may just indicate differences between the two dose computation algorithms' treatment of bone — differences which have been demonstrated previously for other TPSs.<sup>38,39</sup> PerFraction's log-file analysis thus shows promise as a QA tool to detect plans that cannot be delivered accurately, even for complicated treatments such as SBRT spine.

No statistically significant trends could be found between treatment delivery accuracy and MUs, which is in contrast to findings by Kairn et al. where limiting MU substantially increased VMAT plan accuracy and deliverability.<sup>11</sup> This may be due to differences in the optimization algorithms between Eclipse and Pinnacle. In any case, MU cannot be directly controlled in the TPS during inverse planning, so cannot be used as a tool to produce an optimum plan. It was found that plans with greater MU had treatment deliveries with a more negative dose difference at the isocenter. While MU had no impact on plan complexity, there was definite trend with SAS [Fig. 7(d)]. The relationship between MU and SAS makes rational sense: if the prescription is the same and the target volumes are comparable, plans with greater MUs must employ smaller fields. The lack of correlation between MUs and control points means MUs cannot be indirectly influenced by planners in the TPS.

Increasing from 180 to 240 control points (from 4° to 3° between control points), produced consistently better results [Fig. 6(d)] in agreement with the literature.<sup>10</sup> It was shown previously that

90 control point plans were the most complex, and while this had no impact on the plan coverage/conformity, it did have an impact on accuracy/deliverability. Figures 7(a) and 7(b) show that as plan complexity increases, gamma pass rates become consistently worse. This trend has been shown in other studies also<sup>11,12,16</sup> and is most likely due to MLC leaf position discrepancies becoming more frequent<sup>37</sup> rather than model deficiencies.

As expected, plan complexity had no statistically significant impact on isocenter point dose measurements, as dose point placement followed the principles of ICRU 50 to ensure accuracy.<sup>40</sup> SAS had no effect on the gamma pass rates in contrast to the previous studies concerning IMRT plans in Eclipse.<sup>16,41</sup> This may be due to optimization differences between TPSs or due to the fact that, while a greater SAS corresponded with plans delivered with less dose compared to the TPS (according to the isocenter measurement), the effect was not great enough to be noticed by ArcCheck or film gamma criteria. Small fields may not be modeled perfectly, as is often the case when using a single TPS model for all field sizes,<sup>42</sup> but since the effect of SAS is so small it is not recommended for the physicist to spend too much time tweaking this in the model as an avenue of improving their department's SBRT spine accuracy. This study has shown that if the MLC model parameters have been sufficiently validated, accuracy improvements come from producing less complex plans, which is influenced by control points, a parameter that can be easily changed. The planner should feel confident that a less complex plan does not necessarily mean a plan with poorer coverage or conformity, as demonstrated by the lower complexity and higher conformity of plans with 240 control points.

Bringing all these concepts together, only one plan failed ArcCheck and film QA: L2 90X (Fig. 5). This plan had the lowest MCS (0.232) and the highest MI (0.0172), indicating that it was by far the most complex plan produced by the TPS in this study. It was then the most likely to have MLC leaf discrepancies, which is further corroborated by PerFraction showing the lowest pass



rate of all 32 plans (82.8%). Through accuracy optimizations of the SBRT spine treatment technique, one could detect such a plan even before treatment QA or, better yet, avoid producing such a plan altogether.

## 5 | CONCLUSION

A planning and QA study of 32 SBRT spine plans, optimized on four vertebral sites, was conducted in order to find optimum conditions for plan conformity and accuracy of delivery. Plan complexity metrics were found to be independent of plan quality, while QA through four independent methods demonstrated that plan conformity was influenced by the number of control points. In contrast to other studies, it was not MU that improved delivery accuracy but, rather, control points. Like studies on Varian/Eclipse systems, more complex plans were less accurate to deliver, which is attributed to MLC leaf position discrepancies, not solely small fields. The optimum SBRT spine plans in terms of both plan conformity and coverage as well as plan delivery accuracy came from increasing control points to 240.

## CONFLICTS OF INTEREST

The authors have no other relevant conflicts of interest to disclose.

## REFERENCES

1. Timmerman RD. An overview of hypofractionation and introduction to this issue of seminars in radiation oncology. *Semin Radiat Oncol*. 2008;18:215–222.
2. Shaw E, Scott C, Souhami L, et al. Single dose radiosurgical treatment of recurrent previously irradiated primary brain tumors and brain metastases: final report of RTOG protocol 90-05. *Int J Radiat Oncol Biol Phys*. 2000;47:291–298.
3. Cox BW, Spratt DE, Lovelock M, et al. International spine radio-surgery consortium consensus guidelines for target volume definition in spinal stereotactic radiosurgery. *Int J Radiat Oncol Biol Phys*. 2012;83:e597–e605.
4. Ryu S, Fang yin F, Rock J, et al. Image-guided and intensity-modulated radiosurgery for patients with spinal metastasis. *Cancer*. 2003;97:2013–2018.
5. Bilsky MH, Yamada Y, Yenice KM, et al. Intensity-modulated stereotactic radiotherapy of paraspinal tumors: a preliminary report. *Neurosurgery*. 2004;54:823–830.
6. Foote M, Letourneau D, Hyde D, et al. Technique for stereotactic body radiotherapy for spinal metastases. *J Clin Neurosci*. 2011;18:276–279.
7. Wu QJ, Yoo S, Kirkpatrick JP, Thongphiew D, Yin FF. Volumetric arc intensity-modulated therapy for spine body radiotherapy: comparison with static intensity-modulated treatment. *Int J Radiat Oncol Biol Phys*. 2009;75:1596–1604.
8. Dobler B, Khemissi A, Obermeier T, Hautmann MG, Katsilieri Z, Kölbl O. Re-irradiating spinal column metastases using IMRT and VMAT with and without flattening filter - a treatment planning study. *Radiat Oncol*. 2016;11:33.
9. Ghahremani S, Chavez R, Li Y, et al. SU-E-T-435: flattening filter free beams for head and neck IMRT and VMAT optimization. *Med Phys*. 2015;42:3434–3434.
10. Arumugam S, Berry M, Ochoa C, Xing A, Beeksa B, Vial P. Dose discrepancy between planning system estimation and measurement in spine stereotactic body radiation therapy: a case report. *J Med Imaging Radiat Oncol*. 2017;61:695–698.
11. Kairn T, Papworth D, Crowe SB, Anderson J, Christie DR. Dosimetric quality, accuracy, and deliverability of modulated radiotherapy treatments for spinal metastases. *Med Dosim*. 2016;41:258–266.
12. Younge KC, Kuchta JR, Mikell JK, Rosen B, Bredfeldt JS, Matuszak MM. The impact of a high-definition multileaf collimator for spine SBRT. *J Appl Clin Med Phys*. 2017;18:97–103.
13. Ryu S, Pugh SL, Gerszten PC, et al. 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–S132.
14. Webb S. Use of a quantitative index of beam modulation to characterize dose conformality: illustration by a comparison of full beamlet IMRT, few-segment IMRT (fsIMRT) and conformal unmodulated radiotherapy. *Phys Med Biol*. 2003;48:2051–2062.
15. Mcniven AL, Sharpe MB, Purdie TG. A new metric for assessing IMRT modulation complexity and plan deliverability. *Med Phys*. 2010;37:505–515.
16. Crowe SB, Kairn T, Kenny J, et al. Treatment plan complexity metrics for predicting IMRT pre-treatment quality assurance results. *Australas Phys Eng Sci Med*. 2014;37:475–482.
17. Suppli MH, Munck aRP, Pappot H, et al. Stereotactic radiosurgery versus decompressive surgery followed by postoperative radiotherapy for metastatic spinal cord compression (STEREOCORD): study protocol of a randomized non-inferiority trial. *J Radiosur SBRT*. 2016;4:S1.
18. Paddick I. A simple scoring ratio to index the conformity of radiosurgical treatment plans. Technical Note. *J Neurosurg*. 2000;93(Suppl 3):219–222.
19. ICRU. ICRU report 91, prescribing, recording, and reporting of stereotactic treatments with small photon beams. *J ICRU*. 2017; 14:1–160.
20. Sahgal A, Ma L, Gibbs I, et al. Spinal cord tolerance for stereotactic body radiotherapy. *Int J Radiat Oncol Biol Phys*. 2010;77:548–553.
21. Benedict SH, Yenice KM, Followill D, et al. Stereotactic body radiation therapy: the report of AAPM Task Group 101. *Med Phys*. 2010;37:4078–4101.
22. Micke A, Lewis DF, Yu X. Multichannel film dosimetry with nonuniformity correction. *Med Phys*. 2011;38:2523–2534.
23. Hussein M, Clark CH, Nisbet A. Challenges in calculation of the gamma index in radiotherapy - Towards good practice. *Phys Med*. 2017;36:1–11.
24. Pearson K. Note on regression and inheritance in the case of two parents. *Proc Royal Soc London*. 1895;58:240–242.
25. Quan EM, Li X, Li Y, et al. A comprehensive comparison of IMRT and VMAT plan quality for prostate cancer treatment. *Int J Radiat Oncol Biol Phys*. 2012;83:1169–1178.
26. Guckenberger M, Richter A, Krieger T, Wilbert J, Baier K, Flentje M. Is a single arc sufficient in volumetric-modulated arc therapy (VMAT) for complex-shaped target volumes? *Radiother Oncol*. 2009;93:259–265.
27. Prendergast BM, Fiveash JB, Popple RA, et al. Flattening filter-free linac improves treatment delivery efficiency in stereotactic body radiation therapy. *J Appl Clin Med Phys*. 2013;14:4126.
28. Lang S, Shrestha B, Graydon S, et al. Clinical application of flattening filter free beams for extracranial stereotactic radiotherapy. *Radiother Oncol*. 2013;106:255–259.
29. Navarra P, Ascolese AM, Mancosu P, et al. Volumetric modulated arc therapy with flattening filter free (FFF) beams for stereotactic body radiation therapy (SBRT) in patients with medically inoperable early stage non small cell lung cancer (NSCLC). *Radiother Oncol*. 2013;107:414–418.

30. Olch A, Zhuang A. SU-C-BRD-06: sensitivity study of an automated system to acquire and analyze EPID exit dose images. *Med Phys*. 2015;42:3193–3193.
31. Dieterich S, Trestrail E, Holt R, et al. SU-E-T-133: assessing IMRT treatment delivery accuracy and consistency on a Varian TrueBeam using the SunNuclear PerFraction EPID dosimetry software. *Med Phys*. 2015;42:3362–3362.
32. Hsieh ES, Hansen KS, Kent MS, Saini S, Dieterich S. Can a commercially available EPID dosimetry system detect small daily patient setup errors for cranial IMRT/SRS? *Pract Radiat Oncol*. 2017;7:e283–e290.
33. Dinesh Kumar M, Thirumavalavan N, Venugopal Krishna D, Babaiah M. QA of intensity-modulated beams using dynamic MLC log files. *J Med Phys*. 2006;31:36–41.
34. Schreibmann E, Dhabaan A, Elder E, Fox T. Patient-specific quality assurance method for VMAT treatment delivery. *Med Phys*. 2009;36:4530–4535.
35. Qian J, Lee L, Liu W, et al. Dose reconstruction for volumetric modulated arc therapy (VMAT) using cone-beam CT and dynamic log files. *Phys Med Biol*. 2010;55:3597–3610.
36. Defoor DL, Vazquez-quino LA, Mavroidis P, Papanikolaou N, Stathakis S. Anatomy-based, patient-specific VMAT QA using EPID or MLC log files. *J Appl Clin Med Phys*. 2015;16:5283.
37. Li JG, Dempsey JF, Ding L, Liu C, Palta JR. Validation of dynamic MLC-controller log files using a two-dimensional diode array. *Med Phys*. 2003;30:799–805.
38. Zhen H, Hrycushko B, Lee H, et al. Dosimetric comparison of Acuros XB with collapsed cone convolution/superposition and anisotropic analytic algorithm for stereotactic ablative radiotherapy of thoracic spinal metastases. *J Appl Clin Med Phys*. 2015;16:181–192.
39. Carrasco P, Jornet N, Duch MA, et al. Comparison of dose calculation algorithms in slab phantoms with cortical bone equivalent heterogeneities. *Med Phys*. 2007;34:3323–3333.
40. Jones D. ICRU report 50 - prescribing, recording and reporting photon beam therapy". *Med Phys*. 1994;21:833–834.
41. Valdes G, Scheuermann R, Hung CY, Olszanski A, Bellerive M, Solberg TD. A mathematical framework for virtual IMRT QA using machine learning. *Med Phys*. 2016;43:4323.
42. Aspradakis M, Byrne J, Palmans H, et al. IPEM Report 103: Small field MV photon dosimetry. Tech Rep; 2010.