Journal of Radiation Research, Vol. 60, No. 5, 2019, pp. 603–611
doi: 10.1093/jrr/rrz033
Advance Access Publication: 31 May 2019

Impact of delivery characteristics on dose delivery accuracy of volumetric modulated arc therapy for different treatment sites

Jiaqi Li[1], Xile Zhang[1], Jun Li[1], Rongtao Jiang[2], Jing Sui[2], Maria F. Chan[3,*] and Ruijie Yang[1,*]

[1]Department of Radiation Oncology, Peking University Third Hospital, Beijing, China
[2]Brainnetome Center & National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, China
[3]Medical Physics Department, Memorial Sloan Kettering Cancer Center, New York, NY, USA
*Corresponding authors: Ruijie Yang, Department of Radiation Oncology, Peking University Third Hospital, 49 North Garden Rd, Haidian District, Beijing, PR China. Tel: +86-010-8226-4926; E-mail: ruijyang@yahoo.com and Maria F. Chan, Department of Medical Physics, Memorial Sloan Kettering Cancer Center, 136 Mountain View Blvd., Basking Ridge, NJ 07920, USA. Tel:+1-908-542-3046; E-mail: chanm@mskcc.org
(Received 27 December 2018; revised 31 March 2019; editorial decision 24 April 2019)

ABSTRACT

This study aimed to investigate the impact of delivery characteristics on the dose delivery accuracy of volumetric modulated arc therapy (VMAT) for different treatment sites. The pretreatment quality assurance (QA) results of 344 VMAT patients diagnosed with gynecological (GYN), head and neck (H&N), rectal or prostate cancer were randomly chosen in this study. Ten metrics reflecting VMAT delivery characteristics were extracted from the QA plans. Compared with GYN and rectal plans, H&N and prostate plans had higher aperture complexity and monitor units (MU), and smaller aperture area. Prostate plans had the smallest aperture area and lowest leaf speed compared with other plans ($P < 0.001$). No differences in gantry speed were found among the four sites. The gamma passing rates (GPRs) of GYN, rectal and H&N plans were inversely associated with union aperture area (UAA) and leaf speed (Pearson's r: −0.39 to −0.68). GPRs of prostate plans were inversely correlated with aperture complexity, MU and small aperture score (SAS) (absolute Pearson's r: 0.34 to 0.49). Significant differences in GPR between high SAS and low SAS subgroups were found only when leaf speed was <0.42 cm s^{-1} ($P < 0.001$). No association of GPR with gantry speed was found in four sites. Leaf speed was more strongly associated with UAA. Aperture complexity and MU were more strongly associated with SAS. VMAT plans from different sites have distinct delivery characteristics. Affecting dose delivery accuracy, leaf speed is the key factor for GYN, rectal and H&N plans, while aperture complexity, MU and small apertures have a higher influence on prostate plans.

Keywords: complexity metrics; volumetric modulated arc therapy; quality assurance; gamma passing rate

INTRODUCTION

Plan qualities for volumetric modulated arc therapy (VMAT) and intensity modulated radiation therapy (IMRT) are generally equivalent and may vary with treatment sites; the principal advantage of VMAT is the shorter delivery time [1–3]. In most cases, VMAT plans also have a lower number of monitor units (MU) compared with IMRT plans, which result in less whole-body leakage radiation [2]. Therefore, VMAT has gradually become a preferred modality in many clinics [4–6]. VMAT plans usually contain highly modulated beams with dosimetric uncertainties [7, 8]. Moreover, the patient-specific quality assurance (QA) is labor intensive and time consuming [9–11]. With limited accessible information, the degree of beam modulation and its potential impact on dose delivery accuracy are difficult to evaluate prior to QA measurements. The number of MU per beam or plan can be used as a simplistic indicator for plan complexity [12]. Considerable research efforts have been devoted to finding new complexity metrics [13–22].

The first aperture-based metric, the modulation complexity score (MCS), was proposed by McNiven *et al.* [13]. They found that the MCS was effective in characterizing the IMRT plan complexity of

different sites and outperformed MU in detecting dosimetrically robust beams. However, no obvious correlation between the MCS and gamma passing rate (GPR) was observed [13]. In contrast, McGarry *et al.* [14] reported that the MCS was significantly correlated to GPR and minimum segment area in prostate IMRT plans; the same trend between the MCS and GPR was also found in head and neck (H&N) and prostate VMAT plans [15, 16]. Edge metric (EM) proposed by Younge *et al.* [17] was used to penalize irregularly shaped apertures in VMAT plans. Apertures generated with edge penalty were larger and regularly shaped; the delivery error of these apertures decreased >10%. Crowe *et al.* [18] reported that GPR of prostate IMRT beams was significantly correlated to mean field area (MFA) and small aperture score (SAS). Interestingly, the correlations with GPR were found to be diminished for MFA and SAS in VMAT beams [19]. Park *et al.* found that GPRs of H&N and prostate VMAT plans were inversely correlated to leaf speed and acceleration [20, 21]. Du *et al.* found that aperture shape irregularity and MU of IMRT beams greatly decreased with minimum segment area increasing from 2 to 4 cm^2 [22].

To date, many complexity metrics have been developed, and their correlations to IMRT and VMAT dose delivery accuracy have been studied. However, each previous study only evaluated single or several complexity metrics alone, and these results were based on various delivery techniques [segmental-MLC (multi-leaf collimater) IMRT, dynamic-MLC IMRT and VMAT], calculation methods (per beam or per plan), Linac, treatment planning system (TPS) and QA devices. It is difficult to compare the results of different studies, investigate the interactions between different metrics and translate them into IMRT/VMAT QA practice. Therefore, there is a need to classify and have a comprehensive analysis of the metrics in a single uniform clinical setting. Systematically assessing the delivery characteristics of plans from different treatment sites using a variety of metrics can provide more insights into how delivery characteristics affect dose delivery accuracy.

Ten metrics that reflect the fundamental aspect of VMAT delivery characteristics, including aperture complexity, MU, aperture area, leaf speed and gantry speed, were selected. In this study, we focused on investigating the delivery characteristics of different treatment sites, the impact of delivery characteristics on dose delivery accuracy and the correlations between different delivery characteristics.

MATERIALS AND METHODS

VMAT plans and patient-specific QA

Three hundred and forty-four VMAT plans with dual-arc and 2° control-point spacing treated for gynecological (GYN, $n = 123$), H&N (n = 102), rectal ($n = 68$) and prostate ($n = 51$) cancer, respectively, were randomly chosen for retrospective analysis. These treatment sites were selected to represent different delivery characteristics of VMAT plans. The prescription dose to the cervix, uterus, parametrial tissues and pelvic nodes for GYN patients was 50.4 Gy (1.8 Gy/28 fractions, fx). For H&N cases, prescription doses of 60.04 Gy (1.82 Gy/33 fx) and 69.96 Gy (2.12 Gy/33 fx) were delivered to planning target volume (PTV) and planning gross target volume (PGTV), respectively. The prescription dose to gross rectal or rectosigmoid tumor and pelvic nodes for rectal cancer patients was 50.4 Gy (1.8 Gy/28 fx). The prescription dose to prostate and seminal vesicles for prostate cancer patients without lymph node irradiation was 66.0 Gy (2.0 Gy/33 fx). An anisotropic analytical algorithm (AAA) was used for VMAT plan dose calculation; the grid size was 2 mm.

The QA measurement was performed with a MatriXX ion chamber array together with a Multicube phantom (IBA Dosimetry, Schwarzenbruck, Germany). The plan was delivered with Trilogy and Millennium 120 MLC (Varian Medical System, Palo Alto, CA, USA) using the true composite (TC) method as recommended by the AAPM TG-218 report [23]. The angular dependence of the array detector was corrected using a gantry angle sensor during measurement. The absolute dose calibration of the MatriXX ion chamber array was performed before each QA measurement, accounting for Linac daily output fluctuation. The dose profiles of a 20 cm × 20 cm square field were used for array set-up verifications and corrections. Previous studies reported that 2%/2 mm GPR was more sensitive in detecting delivery errors [24–26]; thus, the results of this study were calculated based on 2%/2 mm with 10% dose threshold, absolute dose mode and global normalization (note that 3%/2 mm was used at the time of QA following the recommendation from the TG-218 report).

To extract MLC leaf positions and MU weights from each control point, DICOM RT plans were exported from Eclipse TPS version 10.0 (Varian Medical System) and converted into ASCII format. Then, an in-house-developed MATLAB script was used to calculate the metrics on a per plan basis.

Aperture complexity metrics

The MCS [13] combined leaf sequence variability and aperture shape modulation into one metric. The MCS has a limited scale where 0 is the most complex and 1 is the least complex. EM [17] calculated the ratio of the aperture perimeter defined by the MLC leaf sides to the aperture area. The aperture shape irregularity increases with increased EM.

Plan-normalized monitor units (PMU)

PMU [22] were computed by dividing the total MU of VMAT plans by the fractional target dose and then multiplying by 2 Gy. This parameter is to compare the total MU among VMAT plans with different prescription dose levels.

Aperture area metrics

The MFA [18, 19] was calculated by averaging the area of all individual apertures in a VMAT plan, each aperture area weighted by the number of MU delivered. Union aperture area (UAA) [22] was the union area of all apertures in each arc and was weighted by the number of MU delivered. SAS [18, 19] was used to calculate the proportions of apertures defined as small where the MLC leaf separation was less than a certain value (5, 10 and 20 mm).

Leaf speed metrics

Leaf speed for individual MLC leaves in each control point was calculated by dividing leaf travel distance by delivery time [20, 21].

The average leaf speed (ALS) and proportion of leaf speed ranging from a to b cm s^{-1} [S (a–b)] were calculated. The combinations of a and b were from 0 to 0.4, 0.4 to 0.8, 0.8 to 1.2, 1.2 to 1.6 and 1.6 to 2.0 cm s^{-1}.

Gantry speed metrics

Gantry speed in each control point was calculated by dividing control point spacing by delivery time. The average gantry speed (AGS) and standard deviation of the gantry speed (SGS) were calculated.

Statistics

All statistical analyses were performed with the SPSS statistical software, version 20.0 (IBM Corp., Armonk, NY, USA). To examine the statistical significance of the difference in GPR and metrics between plans from treatment sites, Mann–Whitney U-test and Student's *t*-test were performed, respectively. Pearson correlation analysis, linear regression and univariate analysis were performed to examine the impact of metrics on GPR. Further correlation analysis was performed between different metrics. Statistical significance was defined at two-sided $P < 0.05$. Bonferroni correction was used for multiple comparisons.

RESULTS

Differences in dose delivery accuracy and delivery characteristics among different sites

The 3%/2 mm and 2%2 mm GPR of all sites are shown in Fig. 1. Rectal and H&N plans had the highest and lowest GPR, respectively, while the differences between GYN and prostate plans were not significant. Compared with GYN and rectal plans, H&N and prostate plans generally had higher aperture complexity (lower MCS and higher EM), higher PMU and smaller aperture area (smaller MFA and UAA, and higher SAS) (Fig. 2A–E). Prostate plans had the smallest aperture area and lowest leaf speed [lowest ALS, S (0.4–2.0) and highest S (0–0.4)] than other plans (Fig. 2F, G). GYN and rectal plans shared similar delivery characteristics, which were relatively lower aperture complexity, larger aperture area and higher leaf speed. While both H&N and prostate plans had relatively higher aperture complexity and smaller aperture area, prostate plans had significantly lower leaf speed compared with H&N plans.

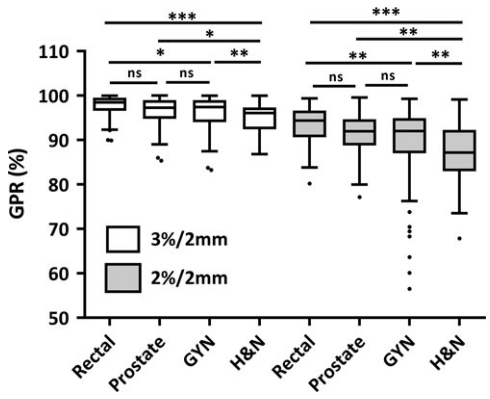

Fig. 1. Gamma passing rate (GPR) for different treatment sites. Results are plotted as Tukey boxplots; dots not included between whiskers are statistical outliers. Mann–Whitney U-test with Bonferroni correction was performed. *$P <$ 0.0083, **$P < 0.0017$, ***$P < 0.00017$. ns = not significant; GYN = gynecological; H&N = head and neck.

In this study, >90% of plans had a constant gantry speed of 4.8° s^{-1} in all control points. Only two GYN, two rectal and three prostate plans had decreased gantry speed; the average number of control points with decreased gantry speed was four (range: 2–7). Twenty-two H&N plans had decreased gantry speed; the average number of control points with decreased gantry speed was 11.18 (range: 4–42). No significant differences in AGS and SGS were found among four treatment sites.

Impact of delivery characteristics on dose delivery accuracy

Two distinct patterns of the correlations between GPR and metrics were observed (Fig. 3). GPR of prostate plans had significant positive correlations with the MCS and MFA (Pearson's *r*: 0.49 and 0.27, respectively) and inverse correlations with EM, PMU and SAS 20 mm (Pearson's *r*: −0.40, −0.41 and −0.34, respectively). No correlations between GPR of prostate plans and UAA, and leaf speed-related metrics were found (Pearson's *r*: −0.03 to 0.08). In contrast, GPR of GYN, rectal and H&N plans had significant positive correlations with S (0–0.4) (Pearson's *r*: 0.53 to 0.70) and inverse correlations with UAA, ALS and S (1.6–2.0) (Pearson's *r*: −0.39 to −0.68). Only weak correlations were found between MCS, EM, PMU, SAS 20 mm and GPR of GYN, rectal and H&N plans (Pearson's *r*: −0.18 to 0.15). The significances of linear relationships between metrics and GPR are shown in Table 1. No correlations between gantry speed and GPR were found. These results showed that leaf speed was the key factor affecting dose delivery accuracy of plans with higher leaf speed (GYN, rectal and H&N plans). Aperture complexity, PMU and SAS were more correlated to dose delivery accuracy of plans with lower leaf speed (prostate plans).

The impact of small apertures on dose delivery accuracy under different leaf speed levels is shown in Fig. 4. All plans from different treatment sites were pooled and stratified into four groups by ALS; each group was then divided into high and low SAS subgroups by the median value of SAS 20 mm. Significant differences in GPR between high SAS and low SAS subgroups were found only when ALS was lower than the first quartile (0.42 cm s^{-1}) (median GPR: 92.45% and 96.83%, respectively, $P < 0.001$). These results suggested that the effect of small apertures on dose delivery accuracy of VMAT was leaf speed dependent.

Correlations between different delivery characteristics

Correlations between aperture area and leaf speed are shown in Fig. 5A. In general, a larger aperture area leads to higher leaf speed [higher ALS, S (1.6–2.0) and lower S (0–0.4)]; UAA played a more important role in affecting leaf speed than SAS and MFA. Aperture

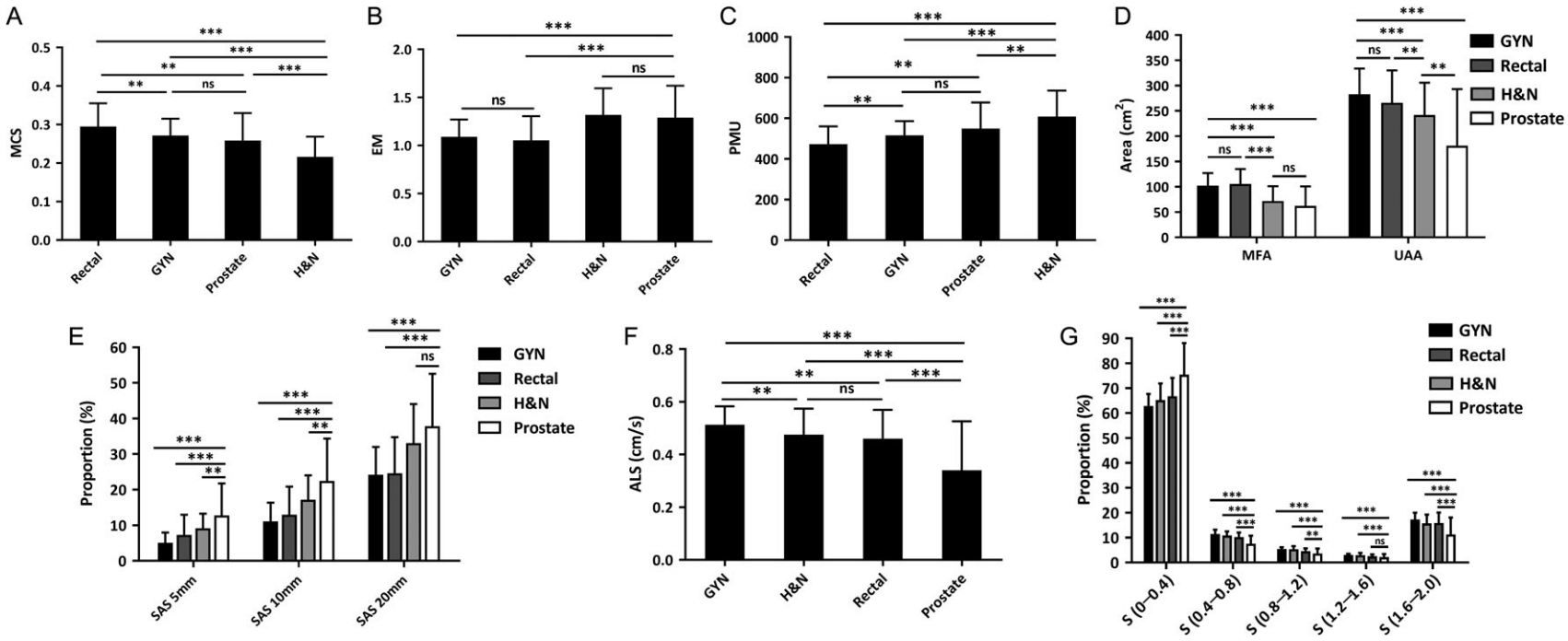

Fig. 2. Differences in delivery characteristics among VMAT plans from different sites. Student's *t*-test with Bonferroni correction was performed. *$P < 0.0083$, **$P < 0.0017$, ***$P < 0.00017$. ns = not significant; GYN = gynecological; H&N = head and neck; MCS = modulation complexity score; EM = edge metric; PMU = plan-normalized MU; SAS = small aperture score; MFA = mean field area; UAA = union aperture area; ALS = average leaf speed; S (a–b) = proportion of leaf speed ranging from a to b cm s^{-1}.

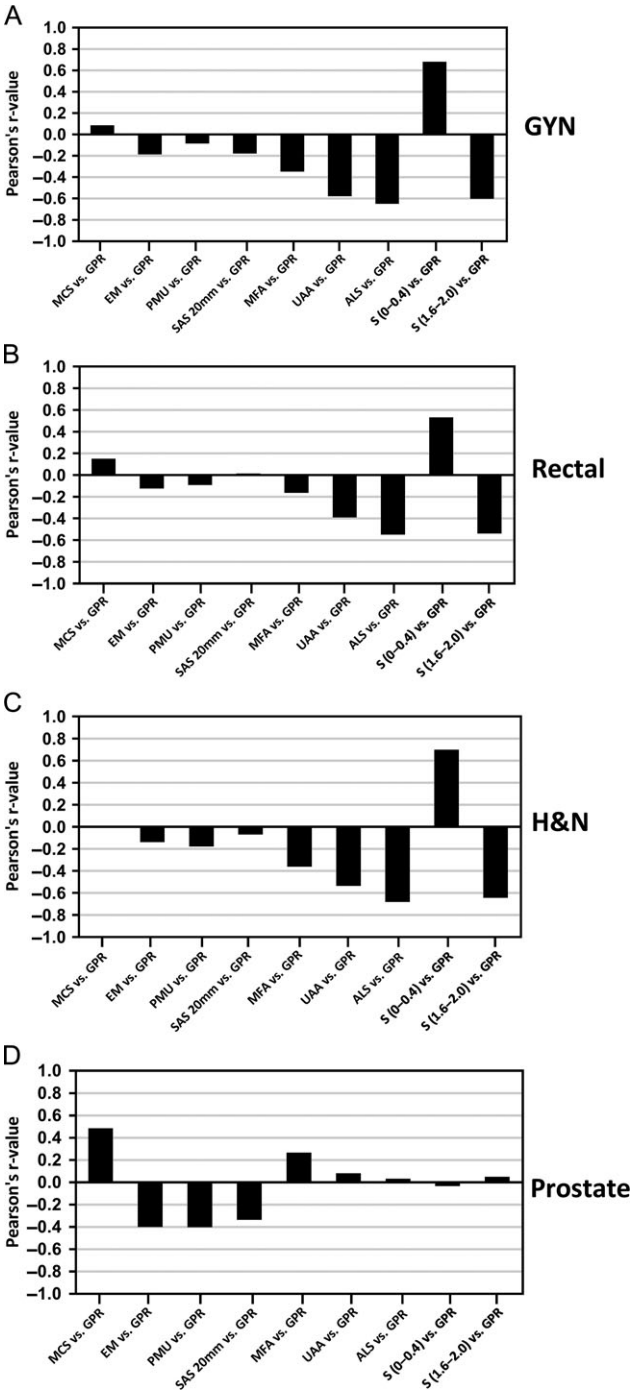

Fig. 3. The correlations between 2%/2 Gamma passing rate (GPR) and metrics in GYN (A), rectal (B), H&N (C) and prostate (D) plans; GYN = gynecological; H&N = head and neck; MCS = modulation complexity score; PMU = plan-normalized MU; SAS = small aperture score; EM = edge metric; MFA = mean field area; UAA = union aperture area; ALS = average leaf speed; S (a–b) = proportion of leaf speed ranging from a to b cm s^{-1}.

complexity increased (lower MCS and higher EM) with increasing SAS and decreasing MFA (Fig. 5B); however, the data did not support a strong role for UAA and leaf speed in influencing aperture complexity. PMU were more correlated to aperture complexity, SAS and MFA than UAA and leaf speed (absolute Pearson's *r*: 0.34–0.85 vs 0.07–0.48) (Fig. 5C). The treatment site dependency was not found in correlations between different delivery characteristics.

DISCUSSION

VMAT plans of different treatment sites have inherent modulation complexity based on target shape and size, location with respect to organ at risk (OAR), prescription dose and OAR dose limits. Our data showed that H&N and prostate plans had higher aperture complexity (lower MCS and higher EM) and PMU due to their smaller MFA and higher SAS compared with GYN and rectal plans. Previous studies [14, 17, 22] also reported that the complexity and MU of IMRT and VMAT plans decreased when the minimum segment area increased. Therefore, to reduce plan complexity and improve delivery efficiency, penalties for small MFA and high SAS could be incorporated into the plan optimization algorithms.

Other major aspects of VMAT delivery characteristics are MLC movement and gantry rotation. By analyzing the log files, many researchers [15, 20, 21, 27] reported that leaf speed was significantly linked to the leaf positional errors in VMAT plans. Kerns *et al.* [27] suggested that restricting the maximum MLC speed could help improve MLC performance for VMAT deliveries. We found that the majority of leaf speed ranged from 0 to 0.4 cm s^{-1} in VMAT plans and the proportion of leaf speed >0.4 cm s^{-1} was a predictor for less dose delivery accuracy for GYN, rectal and H&N plans. Furthermore, we found that larger UAAs were associated with decreased dosimetric accuracy in GYN, rectal and H&N plans. This seems to be counter-intuitive because it is widely believed that a smaller aperture area usually has poorer delivery accuracy. Further correlation analysis from our study has shown that UAAs were positively correlated to leaf speed, and leaf speed was more correlated to UAA than MFA and SAS. This shows that the larger UAA leads to higher leaf speed and eventually leads to lower GPR. In contrast, gantry speed variations were only found in 29 plans and existed in a low proportion of control points; the effect of gantry speed variations on VMAT delivery accuracy was insignificant in this study.

Previous studies [8, 28] reported dose differences up to 30–50% for small static MLC apertures and 5–10% for small volume targets in clinical VMAT plans. Our results demonstrate that the increasing aperture complexity and number of small apertures in prostate plans can lead to noticeable dosimetric errors. H&N plans had comparable aperture complexity and small apertures; however, the dosimetric errors caused by small apertures are insignificant. Moreover, the dosimetric errors caused by small aperture were only significant when ALS was lower than the first quartile (0.42 cm s^{-1}).

In this study, we did not set threshold values for each metric since it is not feasible to precisely predict GPR based on threshold values of single or several complexity metrics; moreover, these threshold values are strictly institution specific. The methodology and metrics in this study can be used for building knowledge-based

Table 1. Linear relationships of different metrics to 2%/2 mm GPR

	GYN		Rectal		H&N		Prostate	
Metrics	R^2 (%)	*P*	R^2 (%)	*P*	R^2 (%)	*P*	R^2 (%)	*P*
MCS	0.72	ns	2.29	ns	0.01	ns	23.62	<0.001
EM	3.49	0.038	1.52	ns	1.92	ns	15.92	0.004
PMU	0.70	ns	0.83	ns	3.17	ns	16.22	0.003
SAS 20 mm	3.24	0.046	0.01	ns	0.51	ns	11.30	0.016
MFA	12.05	<0.001	2.68	ns	13.11	<0.001	7.06	ns
UAA	33.66	<0.001	15.33	0.001	28.57	<0.001	0.66	ns
ALS	42.13	<0.001	30.01	<0.001	46.28	<0.001	0.10	ns
S (0–0.4)	46.31	<0.001	28.33	<0.001	48.88	<0.001	0.15	ns
S (1.6–2.0)	36.31	<0.001	29.10	<0.001	41.76	<0.001	0.25	ns

GPR = gamma passing rate; GYN = gynecological; H&N = head and neck; R^2 = coefficient of determination; ns = not significant; MCS = modulation complexity score; EM = edge metric; PMU = plan-normalized MU; SAS = small aperture score; MFA = mean field area; UAA = union aperture area; ALS = average leaf speed; S (a–b) = proportion of leaf speed ranging from a to b cm s^{-1}; R^2 indicates the percentage variation in GPR explained by each metric.

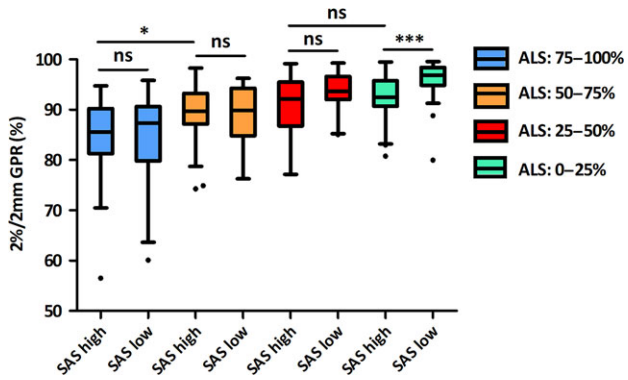

Fig. 4. The impact of average leaf speed (ALS) and small aperture score (SAS) on dose delivery accuracy of VMAT plans. All plans were stratified into four groups by ALS; each group was then divided into high and low subgroups by the median value of SAS. Student's *t*-test with Bonferroni correction was performed. *$P < 0.0083$, **$P < 0.0017$, ***$P < 0.00017$. ns = not significant.

treatment site-specific statistics and assessing new treatment plans with typical complexity levels. For example, if the VMAT QA plan had a higher SAS and lower leaf speed compared with those from prior plans, the physicist should consider checking the suitability of the current calculation model [29]. Studies have shown that Monte Carlo and Acuros XB algorithms had superior dose calculation accuracy for small field compared with AAA and pencil beam convolution (PBC) [30, 31]. Also, physicists should check the accuracy of the percentage depth dose (PDD), profiles and output factor used for small field dosimetry modeling in their TPS [32]. Nelms *et al.* [25] reported that ion chambers with a large sensitive volume (such as a Farmer chamber) have a volume-averaging effect for measurement of small field or high dose gradient beam profiles. Using ion chambers with a small sensitive volume or film dosimetry can improve the accuracy of small field dosimetry modeling [33]. If the VMAT QA plan had a higher ALS or S (0.4–2.0) compared with those from previous plans, stricter or more frequent MLC leaf speed QA could be helpful in preventing delivery errors, such as testing the movement of MLC under various leaf speed levels (i.e. 0.8, 1.2, 1.6 and 2.0 cm s^{-1}) periodically. The knowledge is extremely useful for QA data interpretation, rival plan selection and early detection of 'abnormal' plans. Plans and metrics in this study can also be used to develop and train machine learning models to accurately predict GPR for IMRT [34–37] and potentially VMAT plans.

It is worth noting that dosimetric leaf gap (DLG) also impacts dose calculation accuracy of VMAT. Szpala *et al.* [38] reported that the optimal DLG for VMAT plans was plan specific (range: 0.8–2.0 mm) due to different patterns of leaf movement. Kumaraswamy *et al.* [39] found that for Varian Millennium MLC, due to higher intraleaf transmission of inner 0.5 cm leaf pairs, the optimal DLG for outer 1.0 cm leaf pairs is much lower than DLG values of inner 0.5 cm leaf pairs (average difference: 0.32 mm). The dose differences caused by inaccurate DLG setting become even larger for narrow and small apertures delimited by outer 1.0 cm leaf pairs when MLC gaps are <3.0 cm. In TPS beam modeling, if only the average DLG value for all MLC leaves was used for dose calculation for VMAT plans from different treatment sites, a single and suboptimal DLG value may lead to dose calculation errors. Dose calculation accuracy of VMAT may be improved if the TPS was able to dynamically adjust the DLG value according to the shape and area of MLC apertures. The effects of DLG on GPR were partly reflected by metrics such as EM, MFA and SAS in this study

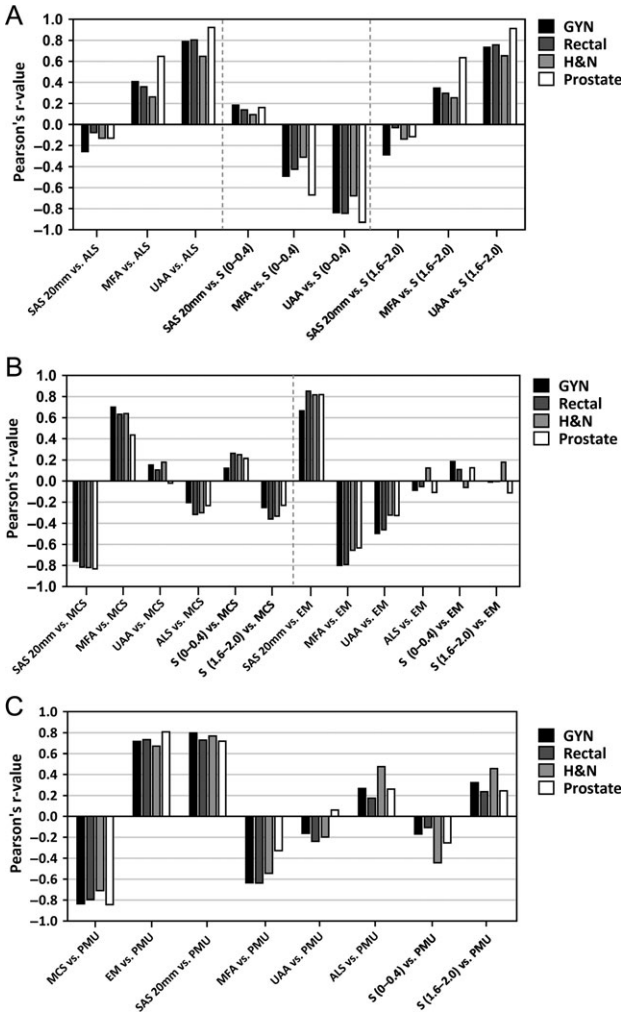

Fig. 5. The correlations between aperture area and leaf speed (A), between aperture complexity, aperture area and leaf speed (B), and between PMU and other metrics (C). GYN = gynecological; H&N = head and neck; MCS = modulation complexity score; EM = edge metric; PMU = plan-normalized MU; SAS = small aperture score; MFA = mean field area; UAA = union aperture area; ALS = average leaf speed; S (a–b) = proportion of leaf speed ranging from a to b cm s^{-1}.

because both the dose calculation errors caused by the improper DLG setting and machine delivery errors caused by MLC mispositioning were more significant in narrow and small apertures.

To the best of our knowledge, this is among the first comprehensive studies of the impact of leaf speed and aperture area on dose delivery accuracy of VMAT for different treatment sites. The evaluation was based on a retrospective setting with a relatively large sample size ($n = 344$), and all treatment plans and QA data were generated from the same TPS and delivered by the same Linac with the same 2D detector array in our department. Previous reports showed that GPR for the same plan varied with the types of dosimeters and QA software [40, 41]. Although not in true 3D with high spatial resolution, dose distributions measured with nonplanar detector arrays can give a more realistic picture of VMAT delivery compared with planar dose distributions using 2D arrays [42–44]. The findings of this study warrant further investigation if using 3D dosimeters.

There have been numerous studies that led to the direction of passing rates correlated with the beam inherent from the IMRT/VMAT plans [13–22, 34–37]. Furthermore, Ford *et al.* [45] conducted a study on quality control quantification from a large database a few years ago and they found that pretreatment IMRT QA is the least effective check among all quality controls for detecting high severity incidents. Fully understanding and dissecting the impact factors of the VMAT delivery accuracy is extremely important for clinical physicists to implement the risk-based program from the AAPM TG-100 report [46]. Thereby, we could save medical physics resources and re-focus our attention on more technical and physics aspects in clinics.

CONCLUSIONS

VMAT plans from different sites have different delivery characteristics. UAA and leaf speed are the key factors affecting the dose delivery accuracy of GYN, rectal and H&N plans, while aperture complexity, PMU and small apertures have a higher impact on the dose delivery accuracy of prostate plans. A larger aperture area leads to higher leaf speed and lower aperture complexity and MU. Moreover, the effect of small apertures on VMAT dose delivery accuracy is dependent on leaf speed.

ACKNOWLEDGEMENTS

This study was presented at the 2018 Chinese Annual Meeting of Radiation Oncology Physics, 20–22 September 2018, Fu Zhou, China.

CONFLICT OF INTEREST

The authors state that they have no conflict of interest with respect to this study

FUNDING

This work was supported by the National Natural Science Foundation of China (grant no. 81071237), Interdisciplinary Medicine Seed Fund of Peking University (grant no. BMU20160585) and the P30 Cancer Center Support Grant from the National Cancer Institute at National Institutes of Health (grant no. CA008748).

REFERENCES

1. Eldebawy E, Parker W, Abdel Rahman W et al. Dosimetric study of current treatment options for radiotherapy in retinoblastoma. *Int J Radiat Oncol Biol Phys* 2012;82:e501–e505.
2. 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;83:1169–78.
3. Clemente S, Wu B, Sanguineti G et al. SmartArc-based volumetric modulated arc therapy for oropharyngeal cancer: a

dosimetric comparison with both intensity-modulated radiation therapy and helical tomotherapy. *Int J Radiat Oncol Biol Phys* 2011;80:1248–55.
4. Abolaban F, Zaman S, Cashmore J et al. Changes in patterns of intensity-modulated radiotherapy verification and quality assurance in the UK. *Clin Oncol (R Coll Radiol)* 2016;28:e28–e34.
5. Barber J, Vial P, White P et al. A survey of modulated radiotherapy use in Australia & New Zealand in 2015. *Australas Phys Eng Sci Med* 2017;40:811–22.
6. Frenzel T, Krüll A. The use of IMRT in Germany. *Strahlenther Onkol* 2015;191:821–6.
7. Otto K. Volumetric modulated arc therapy: IMRT in a single gantry arc. *Med Phys* 2008;35:310–7.
8. Ong CL, Cuijpers JP, Senan S et al. Impact of the calculation resolution of AAA for small fields and RapidArc treatment plans. *Med Phys* 2011;38:4471–9.
9. Oliver M, Gagne I, Bush K et al. Clinical significance of multileaf collimator positional errors for volumetric modulated arc therapy. *Radiother Oncol* 2010;97:554–60.
10. Van Esch A, Huyskens DP, Behrens CF et al. Implementing RapidArc into clinical routine: a comprehensive program from machine QA to TPS validation and patient QA. *Med Phys* 2011; 38:5146–66.
11. Mancuso GM, Fontenot JD, Gibbons JP et al. Comparison of action levels for patient-specific quality assurance of intensity modulated radiation therapy and volumetric modulated arc therapy treatments. *Med Phys* 2012;39:4378–85.
12. Craft D, Süss P, Bortfeld T. The tradeoff between treatment plan quality and required number of monitor units in intensity-modulated radiotherapy. *Int J Radiat Oncol Biol Phys* 2007;67: 1596–605.
13. McNiven AL, Sharpe MB, Purdie TG. A new metric for assessing IMRT modulation complexity and plan deliverability. *Med Phys* 2010;37:505–15.
14. McGarry CK, Chinneck CD, O'Toole MM et al. Assessing software upgrades, plan properties and patient geometry using intensity modulated radiation therapy (IMRT) complexity metrics. *Med Phys* 2011;38:2027–34.
15. Agnew CE, Irvine DM, McGarry CK. Correlation of phantom-based and log file patient-specific QA with complexity scores for VMAT. *J Appl Clin Med Phys* 2014;15:204–16.
16. McGarry CK, Agnew CE, Hussein M et al. The role of complexity metrics in a multi-institutional dosimetry audit of VMAT. *Br J Radiol* 2016;89:20150445.
17. Younge KC, Matuszak MM, Moran JM et al. Penalization of aperture complexity in inversely planned volumetric modulated arc therapy. *Med Phys* 2012;39:7160–70.
18. 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;37:475–82.
19. Crowe SB, Kairn T, Middlebrook N et al. Examination of the properties of IMRT and VMAT beams and evaluation against pre-treatment quality assurance results. *Phys Med Biol* 2015;60: 2587–601.
20. Park JM, Park SY, Kim H et al. Modulation indices for volumetric modulated arc therapy. *Phys Med Biol* 2014;59:7315–40.
21. Park JM, Wu HG, Kim JH et al. The effect of MLC speed and acceleration on the plan delivery accuracy of VMAT. *Br J Radiol* 2015;88:20140698.
22. Du W, Cho SH, Zhang X et al. Quantification of beam complexity in intensity-modulated radiation therapy treatment plans. *Med Phys* 2014;41:021716.
23. Miften M, Olch A, Mihailidis D et al. Tolerance limits and methodologies for IMRT measurement-based verification QA: recommendations of AAPM Task Group No. 218. *Med Phys* 2018;45:e53–e83.
24. Heilemann G, Poppe B, Laub W. On the sensitivity of common gamma-index evaluation methods to MLC misalignments in Rapidarc quality assurance. *Med Phys* 2013;40:031702.
25. Nelms BE, Chan MF, Jarry G et al. Evaluating IMRT and VMAT dose delivery accuracy: practical examples of failure to detect systematic errors when applying a commonly used metric and action levels. *Med Phys* 2013;40:111722.
26. Chan MF, Li J, Schupak K et al. Using a novel dose QA tool to quantify the impact of systematic errors otherwise undetected by conventional QA methods: clinical head and neck case studies. *Technol Cancer Res Treat* 2014;13:57–67.
27. Kerns JR, Childress N, Kry SF. A multi-institution evaluation of MLC log files and performance in IMRT delivery. *Radiat Oncol* 2014;9:176–86.
28. Fog LS, Rasmussen JF, Aznar M et al. A closer look at RapidArc® radiosurgery plans using very small fields. *Phys Med Biol* 2011;56:1853–63.
29. Kerns JR, Stingo F, Followill DS et al. Treatment planning system calculation errors are present in most Imaging and Radiation Oncology Core-Houston phantom failures. *Int J Radiat Oncol Biol Phys* 2017;98:1197–1203.
30. Han T, Mikell JK, Salehpour M et al. Dosimetric comparison of Acuros XB deterministic radiation transport method with Monte Carlo and model-based convolution methods in heterogeneous media. *Med Phys* 2011;38:2651–64.
31. Kan MW, Cheung JY, Leung LH et al. The accuracy of dose calculations by anisotropic analytical algorithms for stereotactic radiotherapy in nasopharyngeal carcinoma. *Phys Med Biol* 2011; 56:397–413.
32. Palmans H, Andreo P, Huq MS et al. Dosimetry of small static fields used in external photon beam radiotherapy: summary of TRS-483, the IAEA-AAPM international Code of Practice for reference and relative dose determination. *Med Phys* 2018;45: e1123–e1145.
33. Arnfield MR, Otto K, Aroumougame VR et al. The use of film dosimetry of the penumbra region to improve the accuracy of intensity modulated radiotherapy. *Med Phys* 2005;32:12–8.
34. Valdes G, Scheuermann R, Hung CY et al. A mathematical framework for virtual IMRT QA using machine learning. *Med Phys* 2016;43:4323–34.
35. Valdes G, Chan MF, Lim SB et al. IMRT QA using machine learning: a multi-institutional validation. *J Appl Clin Med Phys* 2017;18:279–84.
36. Interian Y, Rideout V, Kearney VP et al. Deep nets vs expert designed features in medical physics: an IMRT QA case study. *Med Phys* 2018;45:2672–80.

37. Tomori S, Kadoya N, Takayama Y et al. A deep learning-based prediction model for gamma evaluation in patient-specific quality assurance. *Med Phys* 2018;45:4055–65.
38. Szpala S, Cao F, Kohli K. On using the dosimetric leaf gap to model the rounded leaf ends in VMAT/RapidArc plans. *J Appl Clin Med Phys* 2014;15:67–84.
39. Kumaraswamy LK, Schmitt JD, Bailey DW et al. Spatial variation of dosimetric leaf gap and its impact on dose delivery. *Med Phys* 2014;41:111711.
40. Hussein M, Rowshanfarzad P, Ebert MA et al. A comparison of the gamma index analysis in various commercial IMRT/VMAT QA systems. *Radiother Oncol* 2013;109:370–6.
41. Park JM, Kim JI, Park SY et al. Reliability of the gamma index analysis as a verification method of volumetric modulated arc therapy plans. *Radiat Oncol* 2018;13:175–89.
42. Stathakis S, Myers P, Esquivel C et al. Characterization of a novel 2D array dosimeter for patient-specific quality assurance with volumetric arc therapy. *Med Phys* 2013;40:071731.
43. Wu C, Hosier KE, Beck KE et al. On using 3D γ-analysis for IMRT and VMAT pretreatment plan QA. *Med Phys* 2012;39:3051–9.
44. Boggula R, Lorenz F, Mueller L et al. Experimental validation of a commercial 3D dose verification system for intensity-modulated arc therapies. *Phys Med Biol* 2010;55:5619–33.
45. Ford EC, Terezakis S, Souranis A et al. Quality control quantification (QCQ): a tool to measure the value of quality control checks in radiation oncology. *Int J Radiat Oncol Biol Phys* 2012;84:e263–9.
46. Huq MS, Fraass BA, Dunscombe PB et al. The report of Task Group 100 of the AAPM: application of risk analysis methods to radiation therapy quality management. *Med Phys* 2016;43:4209–62.