

Edge area metric complexity scoring of volumetric modulated arc therapy plans

Julia Götstedt^{a,*}, Anna Bäck^{a,b}

^a Department of Radiation Physics, Institute of Clinical Sciences, Sahlgrenska Academy, University of Gothenburg, 413 45 Gothenburg, Sweden

^b Department of Therapeutic Radiation Physics, Medical Physics and Biomedical Engineering, Sahlgrenska University Hospital, 413 45 Gothenburg, Sweden

ARTICLE INFO

Keywords:

Complexity
Complexity metric
EAM
Quality
Quality control
VMAT

ABSTRACT

Background and purpose: Aperture-based complexity metrics have been suggested as a method to score complexity of volumetric modulated arc therapy (VMAT) plans. The purpose of this study was to evaluate the edge area metric (EAM) for clinical VMAT plans on a control point and treatment plan level.

Materials and methods: EAM on a control point level was evaluated based on film measurements of 18 static beam openings originating from VMAT plans. EAM on a treatment plan level (arithmetic mean value of EAM scores for control points) was evaluated based on measurements with the Delta⁴® for 200 VMAT plans for four different treatment sites: pelvic, thorax, head and neck, and prostate. Measurements were compared to calculations and dose difference and gamma pass rates were evaluated.

Results: EAM scores on a control point level correlated with Pearson's r-values of -0.96 and -0.77 to dose difference and gamma pass rates, respectively. The prostate plans had the highest average EAM score. A connection between smaller PTVs and higher EAM scores was found. No correlation between the evaluation result and EAM on a plan level was found.

Conclusions: EAM on a control point level was shown to correlate to the difference between measured and calculated 2D dose distributions of clinical VMAT beam openings. No correlation was found for EAM on a plan level for clinical treatment plans.

1. Introduction

Volumetric modulated arc therapy (VMAT) involves simultaneous alteration of the beam opening formed by multi-leaf collimators (MLC), gantry speed, and dose rate, for beam modulation during treatment [1]. A treatment field, i.e. arc, is divided in control points for which the MLC leaf positions, gantry angle, and cumulative number of monitor units (MU) are defined.

Treatment plans with high amount of complex beam openings, i.e. small and/or irregularly shaped, will create relatively large regions that lack charged particle equilibrium where clinical dose calculation algorithms have difficulties to accurately calculate the dose distribution [2]. Furthermore, complex beam openings make the delivered dose more sensitive to small variations in the treatment machine settings, e.g. the positioning of the MLC leaves [3,4]. Therefore, treatment plans with high amount of complex beam openings might lead to clinically relevant dose differences between planned and delivered dose distributions [2–5].

Aperture-based complexity metrics have been suggested to score the complexity of VMAT plans to detect complexity at an early stage of treatment preparation for either re-consideration, re-planning, or extra attention during quality controls (QC) [6,7]. A complexity metric could also be integrated in the objective function of the optimization process in the TPS to automatically make treatment plans less complex [8,9].

Before making clinical decisions based on a complexity metric, thorough evaluation of the metric should be performed. Numerous comparative studies of different complexity metrics and their correlation to measurement results of commonly used pre-treatment QC methods have been performed [6,7,10–15]. However, dedicated evaluations of metrics with the purpose to increase information on the dosimetric uncertainty of a plan beyond common QC results, are scarce.

The edge area metric (EAM) has shown promising results in a previous evaluation study where different complexity metrics were compared based on artificially generated static beam openings [16]. To our knowledge, EAM scores have not yet been evaluated for clinical VMAT beam openings. The purpose of this study was to evaluate EAM

* Corresponding author.

E-mail address: julia.gotstedt@vregion.se (J. Götstedt).

<https://doi.org/10.1016/j.phro.2021.02.002>

Received 25 June 2020; Received in revised form 12 February 2021; Accepted 23 February 2021

Available online 6 March 2021

2405-6316/© 2021 The Authors. Published by Elsevier B.V. on behalf of European Society of Radiotherapy & Oncology. This is an open access article under the

CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

for clinical VMAT plans on a control point and treatment plan level for different treatment sites.

2. Material and method

EAM is based on dividing a beam opening into a complex (R_1) and a non-complex (R_2) region and is defined as $EAM = R_1/(R_1 + R_2)$ [16] (Supplementary Fig. 1). The complex region encloses an area both inside and outside of the beam opening edge and the non-complex region is defined as the rest of the open area within the beam opening. Hence, EAM score is a positive value up to 1 for the most complex openings. Besides the absolute level on the EAM scale, only small relative variations in EAM have been observed when varying the extent of R_1 from 1 mm to 5 mm (at isocenter distance). The largest separation between lowest and highest EAM scores for control points originating from the same treatment was found for R_1 of 2.5 mm which was used in this study. The EAM calculations were performed using an in-house developed MATLAB® software.

VMAT plans used in this study were approved and used for treatments within four groups of treatment sites: larger targets in the pelvic, targets in the thorax and the head and neck (H&N) region, and prostate cancer treatments. This project was classified as a quality development and assurance project without processing of personal data and no statement of access to use the treatment plans was required. The plans were created for Clinac iX or TrueBeam STx (Varian Medical Systems) equipped with Millennium and High Definition MLC respectively. Dose calculations were performed using analytical anisotropic algorithm (AAA) in Eclipse (version 13.6) with calculation grid size 0.25 cm × 0.25 cm. The minimum leaf gap was set to 0.05 cm and a full arc (358 deg) were divided in 178 control points. In general, the treatment plans for prostate cancer and larger targets in the pelvic region included 2–3 full arcs while the plans created for targets in the H&N and thorax region were more diverse but, in most cases included 2–3 either full or partial arcs.

2.1. EAM for VMAT control points

Six full arc treatment plans – two H&N plans (two arcs each), three prostate plans (one single and two double arc plans) and one plan created for treatment of vulvar cancer (3 arcs) – were selected for evaluation of EAM on a control point level. From each plan, beam openings from three control points in one of the arcs were selected

Table 1

Information about the 18 beam openings for evaluation of EAM on a control point level. The beam openings are grouped according to their EAM score. The MLC type is either Millennium (M) or High Definition (HD).

Beam opening number	EAM score	Treatment machine	MLC type	Treatment site	Beam opening area [cm ³]	Number of sub openings
1	0.35	Clinac iX	M-MLC	Pelvic	109.3	2
2	0.35	Clinac iX	M-MLC	H&N	120.8	3
3	0.41	TrueBeam STx	HD-MLC	H&N	85.5	3
4	0.41	TrueBeam STx	HD-MLC	Prostate	26.8	1
5	0.41	TrueBeam STx	HD-MLC	Prostate	29.9	1
6	0.55	Clinac iX	M-MLC	Pelvic	55.0	8
7	0.55	Clinac iX	M-MLC	H&N	63.0	3
8	0.55	Clinac iX	M-MLC	Prostate	24.9	2
9	0.74	TrueBeam STx	HD-MLC	H&N	44.6	8
10	0.74	Clinac iX	M-MLC	Pelvic	57.5	7
11	0.74	Clinac iX	M-MLC	H&N	39.8	6
12	0.74	TrueBeam STx	HD-MLC	Prostate	16.6	2
13	0.87	TrueBeam STx	HD-MLC	H&N	25.7	10
14	0.87	Clinac iX	M-MLC	Prostate	14.5	6
15	0.87	TrueBeam STx	HD-MLC	Prostate	14.2	2
16	0.99	Clinac iX	M-MLC	Prostate	6.4	10
17	0.99	TrueBeam STx	HD-MLC	Prostate	9.0	7
18	0.99	TrueBeam STx	HD-MLC	Prostate	3.7	12

(Table 1). The beam openings were selected to cover the whole range of EAM scores for the plans (0.35 to 0.99). Beam openings with the same EAM score originating from different plans were also selected. Examples of beam openings are given in Supplementary material (Supplementary Fig. 2).

The dose distribution of the beam openings was measured at the respective treatment machine using Gafchromic EBT3 films. The double exposure measurement procedure used, with pre-scan and pre-exposure, is described previously [16]. The measurements were performed with the gantry in 0 degrees and the film at 10 cm depth in a solid water phantom with a source-surface-distance (SSD) of 90 cm and calculated in the same geometry. The number of MU was adjusted to result in a calculated maximum dose of 2 Gy ($\pm 0.2\%$) at 10 cm depth. Each measurement was repeated on three different occasions and the result is reported as the mean value of these measurements to account for variations of sensitivity within and between separate film sheets and other uncertainties related to the measurement and evaluation procedure as well as to account for treatment machine delivery variations.

The measured and calculated dose distributions were imported into the evaluation software RIT113 (Radiological Imaging Technology©). The hair cross from the measurement setup was marked in the margin of the film and used for rotation correction and to define the isocenter in the measured dose distributions. The isocenter was used for registration to the corresponding calculated dose distributions. This alignment method is independent of potential inaccuracies in the dose calculation due to a non-proper modelling of the tongue and groove design of the MLC leaves which would shift the position of the penumbra in a direction opposed to the MLC leaf direction. The median dose value within a centrally placed region of interest (0.5 cm × 0.5 cm or 0.25 cm × 0.25 cm for the three smallest openings) in both measured and calculated dose distributions were used for normalization. The calculated and measured dose distributions were compared pixel-by-pixel with a global dose difference criterion of 5% and with gamma evaluation using 3%, 1 mm criterion. A stricter criterion than 5%, without including a DTA criteria would be in the range of the inherent uncertainty of the measurement procedure. The EAM scores were compared to the relative number of pixels within the evaluation criteria (threshold of 10% of the maximum calculated dose).

2.2. EAM for VMAT plans

The arithmetic mean value of the EAM scores calculated on a control

point level was evaluated for scoring complexity on a treatment plan level. Control points included in avoidance sectors with zero dose rate were excluded.

EAM scores were calculated for 200 of the most recent VMAT plans (50 plans for each of the four treatment sites) and analyzed on both a treatment plan and on a control point level. The design of EAM implicitly means that a beam opening with a smaller open area has a higher EAM and the beam opening area is often related to the projection size of the planning target volume (PTV). To study this relation, the PTV volumes were extracted. In the case of multiple PTVs, the union of all PTV volumes was extracted.

The plans were measured on one occasion before treatment start using the Delta⁴® (ScandiDos) phantom described previously [17–19]. A half arc with a 10 cm × 10 cm beam opening formed using the collimator jaws was also measured to correct for the daily machine output variation. The dose distributions were aligned using the optimized phantom position tool provided in the Delta⁴® software. The calculated and measured dose were compared in the diode measurement points. The pass rate for the evaluated points (receiving higher than 20% of the maximum dose) within a global dose difference of 3% normalized to the maximum dose in the measurements, as well as within a gamma criterion of 3%, 1 mm, were compared to EAM on a plan level. A stricter criterion would be in the range of the dosimetric accuracy [18].

3. Results

3.1. EAM for VMAT control points

The linear correlations, expressed as Pearson's *r*-values, between EAM scores and evaluated differences between measurements and calculations on a control point level were -0.96 and -0.77 for dose difference and gamma pass rates respectively (Fig. 1). The spread in mean dose difference and gamma pass rates for beam openings from different treatment plans but with the same EAM score was larger for more complex beam openings with higher EAM (Fig. 1 and Supplementary Table 1).

3.2. EAM for VMAT plans

The average EAM score for all plans was 0.59 and 0.67, 0.59, 0.58 and 0.52 for prostate plans, thorax plans, H&N plans, and plans planned for larger targets in the pelvic, respectively. The proportion of the prostate plans with an EAM score higher than the total average (0.59) was 94% and the corresponding proportion for the plans planned for larger targets in the pelvic was 6%.

The EAM results per treatment site on a plan level were confirmed by the results of EAM on a control point level (Fig. 2). The higher complexity found for prostate plans was also indicated from the dose difference pass rate analysis shown for comparison (Fig. 2c). On average, the prostate plans had a lower EAM for the lateral beam openings as seen in polar plots for EAM as a function of gantry angle (Supplementary Fig. 3). No such average trends could be seen for the other treatment sites. The group of prostate cancer patients included a larger number of patients with smaller PTVs compared to the other three groups. A connection between smaller PTVs and higher EAM scores was found (Fig. 3).

No linear correlation was found between EAM scores and evaluated differences between measurements and calculations on a treatment plan level (Fig. 4). Two plans, P1 and P2 (Fig. 4), stood out with respect to high EAM scores of 0.81 and 0.80 respectively. P1 was planned for treatment of a small PTV (32 cm³) in the H&N and P2 was planned for a postoperative treatment in the thorax region (PTV 76 cm³). Comparison of EAM on a control point level for P1 and P2 to the other plans of respective treatment site showed that the P1 and P2 histograms were shifted towards higher complexity (Fig. 5). The increased complexity of the P1 plan was found for almost all separate gantry angles when compared to 6 H&N plans planned for the same angles (Fig. 5a).

4. Discussion

The performance of EAM was evaluated on a control point and on a treatment plan level for VMAT plans. A correlation of EAM scores and dose differences between measured and calculated dose distributions was found on control point level. The possibility to evaluate complexity on control point level to provide information of complexity variations as a function of gantry angle can be advantageous in the treatment plan evaluation process. No correlation to differences between calculated and measured dose was found when the mean value of EAM (on control point level) was used to quantify the complexity on treatment plan level. The EAM scores were found to be treatment site dependent and the prostate plans were found to have the highest average EAM.

The measurement depth of 10 cm represents a relevant tumor depth and has been used in other studies [8]. The evaluated pass rates are expected to be different at other depths but the relation to EAM is expected to be similar. However, because dose differences are expected to be more pronounced at shallower depths, evaluations at one single depth is a limiting factor. The calculation grid size has an influence on both the pass rates and the correlation to EAM [20] and the grid size was chosen based on common clinical use. For static fields with the same EAM score, the spread in pass rates were larger for the group of fields

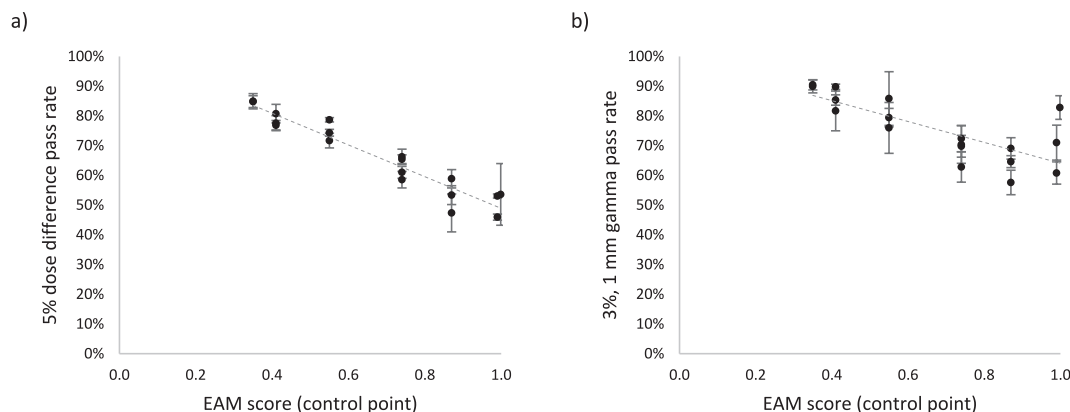


Fig. 1. The calculated EAM complexity scores for the beam openings of the 18 selected VMAT control points versus a) the 5% dose difference pass rate and b) 3%, 1 mm gamma pass rate. The evaluation result for each point is the mean value of three film measurements and the error bars plotted in the figure are the standard deviations from the three measurements. A linear trend line is fitted to the result.

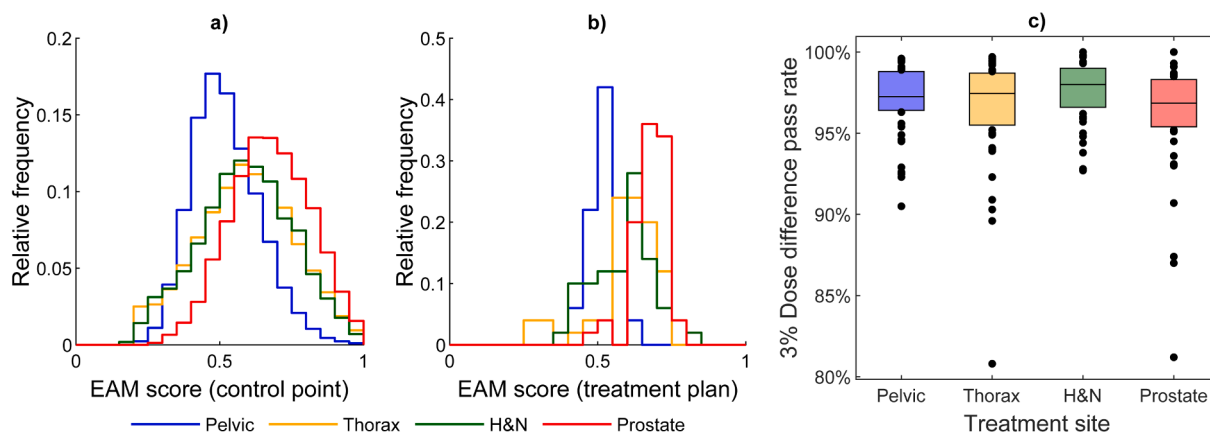


Fig. 2. The complexity analysis of the 200 VMAT plans is illustrated for four treatment sites (50 plans in each group) as a) the relative frequency of EAM scores on control point level, b) the relative frequency of EAM scores on treatment plan level and c) the 3% dose difference pass rate results for the evaluated Delta⁴ measurements and AAA calculations. The area under the curves in the frequency histograms are normalized to 1. The boxes in c) are defined as the data within the 25th and 75th percentiles, and the median values of respective dataset is marked with a horizontal line.

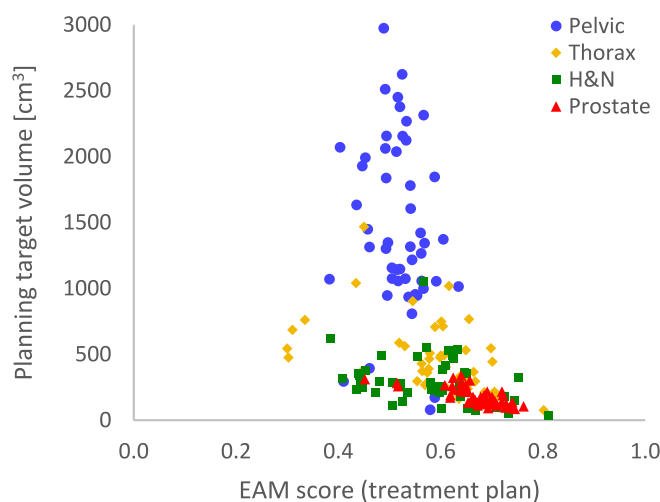


Fig. 3. The planning target volume (PTV) (the union volume when multiple PTVs are defined) for 200 patients and the EAM scores calculated on a treatment plan level for the corresponding VMAT plans studied. The results are separated for the group of patients treated for larger targets in the pelvic region (blue circles), thorax treatments (yellow diamonds), H&N treatments (green squares) and prostate treatments (red triangles). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

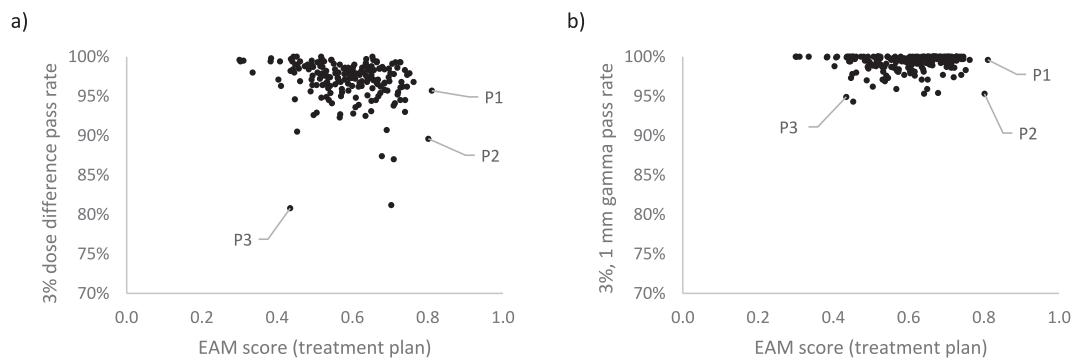


Fig. 4. Scatter plots for the EAM scores on a treatment plan level (mean value of EAM scores on a control point level) and a) the 3% dose difference pass rate and b) 3%, 1 mm gamma pass rate for the evaluated Delta⁴ measurements and AAA calculations. Three plans are identified to be subject for further investigation. Two plans (P1, P2) due to the high EAM scores and one plan (P3) due to the combination of a low dose difference pass rate and low EAM score.

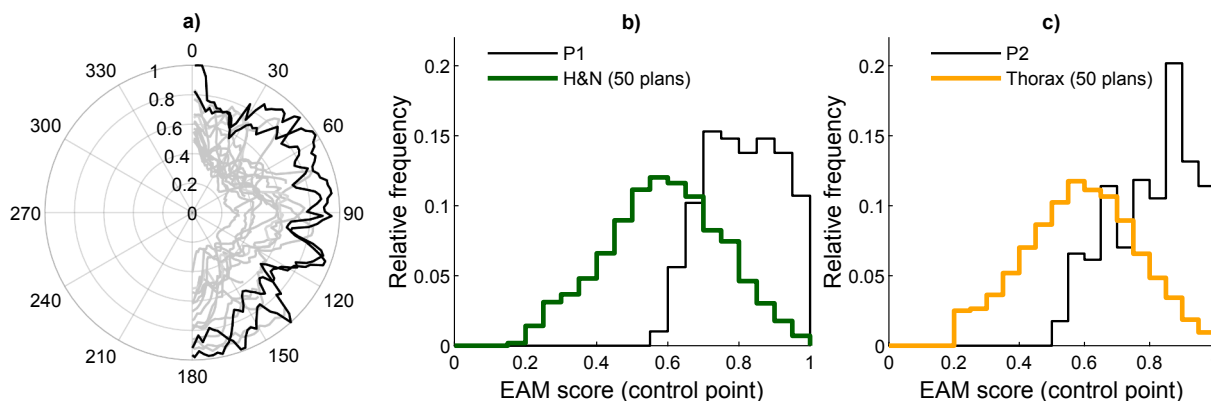


Fig. 5. a) The two arcs of the P1 H&N plan (from Fig. 4) is plotted as a polar plot with the EAM scores on a control point level (r-axis) plotted for corresponding gantry angle (θ -axis) as black lines compared to 6 H&N plans using the same treatment arcs plotted as grey lines. b) The P1 H&N plan is shown in a frequency histogram for EAM scores on a control point level for comparison with the EAM scores on a control point level for all 50 H&N plans. c) The P2 thorax plan for EAM on a control point level is compared to all 50 thorax plans. The area under the curves in the frequency histograms are normalized to 1.

that were ranked as more complex. This can be explained by an increased uncertainty that affects the variations in the delivered dose for different occasions but also the sensitivity to measurement uncertainties. This supports the argument that it is important to decrease the amount of complex beam openings to create high quality and robust treatment plans.

An evaluation method including a DTA criteria, such as the gamma evaluation, can be used to disregard some dose differences due to systematic and random measurement errors. However, a DTA criterion will also disregard true differences between calculated and delivered dose. In fact, commonly used evaluation methods which includes a DTA criterion has been shown to be an insufficient tool for the purpose of detecting differences between calculated and measured dose distributions [6–7,11,21–25]. The evaluation of dose differences in the penumbra region is challenging since a relatively large dose deviation in a specific point could be due to a relatively small alignment error. Evaluations that exclude a DTA criterion therefore need careful alignment methods but will nevertheless lead to larger deviations, especially in high dose regions, which is a limitation. The choice of evaluation criteria for gamma evaluation is critical [25–28] and should be selected based on the specific purpose. The criteria must also be selected pursuant to the inherent characteristics of the calculation and measurement procedures [25,29]. Evaluations based on dose difference alone are expected to result in lower pass rates as compared to gamma evaluations but will also result in a higher differentiation of treatment plans (Fig. 4).

The prostate plans had the highest average EAM score, probably mainly due to the smaller PTV. In our experience, a beam opening originating from a prostate plan can be quite irregular despite the regular form of the target (Supplementary Fig. 2). The complexity is dependent on the number of arcs used in the treatment plans [12] and on the TPS and treatment machine [30]. Thus, a complexity score should be interpreted in relation to a baseline of plans for the same treatment site, technique, TPS, and treatment machine.

One plan (P3) stood out with a low pass rate but at the same time a low EAM (Fig. 4a). This plan was created to treat Hodgkin's lymphoma with three anterior partial VMAT arcs. The calculated dose in the Delta⁴ geometry resulted in several high dose gradient regions which might have led to an increased measurement uncertainty. The PTV (1040 cm³) was relatively large in comparison to the other patients (mean 590 cm³, range 30 to 2970 cm³) which partly explains the low EAM.

EAM has been categorized as an accuracy metric [7]. Other examples of accuracy metrics evaluated for VMAT are based on the opening area alone, “segment area per control point” [12], and “small aperture score” [11], based on the irregularity of the opening alone, “leaf offset impact on calculation” [15], while some, like EAM, combines area and

irregularity of the opening, “converted aperture metric” [16], and “edge metric” [8,13]. The dynamic nature of the VMAT delivery can also lead to an increased uncertainty which is not taken into account by the purely aperture-based accuracy metrics. The “comprehensive modulation index” [31] was categorized as a combination of an accuracy and a deliverability metric.

Even though a correlation between EAM and dosimetric evaluations was found for VMAT control point beam openings it does not directly translate to a correlation on treatment plan level. One reason could be that the mean value blur information detected by EAM on a control point level. It might also be related to limitations in the quasi-3D measurement method used. Lack of correlation on a plan level for accuracy metrics has also been reported by others [10–12], unless evaluating individual structure volumes separately [12]. Our study included exclusively plans that were approved for treatment. Other studies that included both plans that passed and failed the clinical QC were able to show that accuracy metrics could distinguish plans that failed from those that passed [13,15]. All this underlines the importance of defining the purpose of a complexity metric and to validate that the metric meets that specific purpose.

To interpret the clinical relevance of aperture-based complexity, the dose contribution should be considered. Some suggested VMAT complexity metrics include a MU weighting factor [13,14,24,30]. Introducing a weight parameter to decrease the impact of control points with lower MUs raises the question of the most suitable relation between that weight and the number of MUs. For example, a highly complex beam opening needs a higher number of MUs to deliver the same dose compared to a less complex beam opening, but a highly complex beam opening often has a smaller area and consequently contributes to a smaller volume. Furthermore, the clinical impact depends on the combination of the amount of dose and location within the patient. The translation from MU to clinically relevant dose contribution is not straightforward. Interpretation of patient-specific clinical relevance of plan complexity is studied in a parallel work by our group [32].

In conclusion, the EAM has been shown to correlate to differences between measured and calculated dose distributions for clinical VMAT control points and can be used to quantify the variation of complexity in different parts of the treatment arc. The EAM complexity score should be interpreted in relation to the treatment site. The complexity on treatment plan level quantified as a mean value of EAM control point scores was found to not correlate to differences between measured and calculated dose distributions of clinical VMAT plans.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We acknowledge the financial support by the Swedish Radiation Safety Authority, the King Gustaf V Jubilee Clinic Cancer Research Foundation and the Health & Medical Care Committee of the Regional Executive Board, Region Västra Götaland and are thankful for the support from our colleagues at the Department of Therapeutic Radiation Physics.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.phro.2021.02.002>.

References

- [1] Otto K. Volumetric modulated arc therapy: IMRT in a single gantry arc. *Med Phys* 2008;35:310. <https://doi.org/10.1118/1.2818738>.
- [2] Das IJ, Ding GX, Ahnesjo A. Small fields: nonequilibrium radiation dosimetry. *Med Phys* 2008;35:206–15. <https://doi.org/10.1118/1.2815356>.
- [3] LoSasso T, Chui CS, Ling CC. Physical and dosimetric aspects of a multileaf collimation system used in the dynamic mode for implementing intensity modulated radiotherapy. *Med Phys* 1998;25:1919–27. <https://doi.org/10.1118/1.598381>.
- [4] Oliver M, Gagne I, Bush K, Zavgorodni S, Ansbacher W, Beckham W. Clinical significance of multi-leaf collimator positional errors for volumetric modulated arc therapy. *Radiother Oncol* 2010;97:554–60. <https://doi.org/10.1016/j.radonc.2010.06.013>.
- [5] Fog LS, Rasmussen JFB, Aznar M, Kjaer-Kristoffersen F, Vogelius IR, Engelholm SA, et al. A closer look at RapidArc radiosurgery plans using very small fields. *Phys Med Biol* 2011;56:1853–63. <https://doi.org/10.1088/0031-9155/56/6/020>.
- [6] Antoine M, Ralite F, Soustiel C, Marsac T, Sargos P, Cugny A, et al. Use of metrics to quantify IMRT and VMAT treatment plan complexity: a systematic review and perspectives. *Phys Med* 2019;64:98–108. <https://doi.org/10.1016/j.ejmp.2019.05.024>.
- [7] Chiavassa S, Bessieres I, Edouard M, Mathot M, Moignier A. Complexity metrics for IMRT and VMAT plans: a review of current literature and applications. *Br J Radiol* 2019;92:20190270. <https://doi.org/10.1259/bjr.20190270>.
- [8] Younge KC, Matuszak MM, Moran JM, McShan DL, Fraass BA, Roberts DA. Penalization of aperture complexity in inversely planned volumetric modulated arc therapy. *Med Phys* 2012;39:7160–70. <https://doi.org/10.1118/1.4762566>.
- [9] Binny D, Spalding M, Crowe SB, Jolly D, Kairn T, Trapp JV, et al. Investigating the use of aperture shape controller in VMAT treatment deliveries. *Med Dosim* 2020;45:284–92. <https://doi.org/10.1016/j.meddos.2020.02.003>.
- [10] Du W, Cho S, Zhang X, Hoffman K, Kudchadker R. Quantification of beam complexity in intensity-modulated radiation therapy treatment plans. *Med Phys* 2014;41:021716. doi:10.1118/1.4861821.
- [11] Crowe S, Kairn T, Middlebrook N, Sutherland B, Hill B, Kenny J, 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. <https://doi.org/10.1088/0031-9155/60/6/2587>.
- [12] Shen L, Chen S, Zhu X, Han C, Zheng X, Deng Z, et al. Multidimensional correlation among plan complexity, quality and deliverability parameters for volumetric modulated arc therapy using canonical correlation analysis. *J Radiat Res* 2018;59:207–15. <https://doi.org/10.1093/jrr/rrx100>.
- [13] Younge KC, Roberts D, Janes LA, Anderson C, Moran JM, Matuszak MM. Predicting deliverability of volumetric-modulated arc therapy (VMAT) plans using aperture complexity analysis. *J Appl Clin Med Phys* 2016;17:124–31. <https://doi.org/10.1120/jacmp.v17i4.6241>.
- [14] McNiven AL, Sharpe MB, Purdie TG. A new metric for assessing IMRT modulation complexity and plan deliverability. *Med Phys* 2010;37:505–15. <https://doi.org/10.1118/1.3276775>.
- [15] Dechambre D, Baart V, Mathot M. OC-0612: dedicated VMAT complexity metrics reduce patient QA workload. *Radiother Oncol* 2018;127:323. [https://doi.org/10.1016/S0167-8140\(18\)30922-8](https://doi.org/10.1016/S0167-8140(18)30922-8).
- [16] Götstedt J, Karlsson AH, Bäck A. Development and evaluation of aperture-based complexity metrics using film and EPID measurements of static MLC openings. *Med Phys* 2015;42:3911–21. <https://doi.org/10.1118/1.4921733>.
- [17] Feygelman V, Forster K, Opp D, Nilsson G. Evaluation of biplanar diode array dosimeter for quality assurance of step-and-shoot IMRT. *J Appl Clin Med Phys* 2009;10:64–78. <https://doi.org/10.1120/jacmp.v10i4.3080>.
- [18] ScandiDos, Delta4®. Technical Specifications – Delta4 Phantom+, 1 System; 2018.
- [19] Bedford J, Lee Y, Wai P, South C, Warrington A. Evaluation of the Delta⁴ phantom for IMRT and VMAT verification. *Phys Med Biol* 2009;54:167–76. <https://doi.org/10.1088/0031-9155/54/9/N04>.
- [20] Larsson L, Bäck A, Karlsson A. The impact of different dose calculation algorithms and grid sizes on aperture-based complexity metrics. *Dep. Radiation Physics, Gothenburg University*; 2015. https://radfys.gu.se/digitalAssets/1559/1559866_axel_larsson_enskilt-arbete_2015vt.pdf.
- [21] Zhen H, Nelms BE, Tome WA. Moving from gamma passing rates to patient DVH-based QA metrics in pretreatment dose QA. *Med Phys* 2011;38:5477–89. <https://doi.org/10.1118/1.3633904>.
- [22] Nelms BE, Zhen H, Tome WA. Per-beam, planar IMRT QA passing rates do not predict clinically relevant patient dose errors. *Med Phys* 2011;38:1037–44. <https://doi.org/10.1118/1.3544657>.
- [23] Nilsson J, Hauer AK, Bäck A. IMRT patient-specific QA using the Delta⁴ dosimetry system and evaluation based on ICRU 83 recommendations. *J Phys Conf Ser* 2013;444:012048. <https://doi.org/10.1088/1742-6596/444/1/012048>.
- [24] Stasi M, Bresciani S, Miranti A, Maggio A, Sapino V. Pretreatment patient-specific IMRT quality assurance: a correlation study between gamma index and patient clinical dose volume histogram. *Med Phys* 2012;39:7626–7634. doi:10.1118/1.4767763.
- [25] Steers J, Fraass B. IMRT QA: selecting gamma criteria based on error detection sensitivity. *Med Phys* 2016;43:1982–1994. doi:10.1118/1.4943953.
- [26] Nelms BE, Jarry G, Chan M, Hampton C, Watanabe Y, Feygelman V. Real-world examples of sensitivity failures of the 3%/3mm pass rate metric and published action levels when used in IMRT/VMAT system commissioning. *J Phys Conf Ser* 2013;444:012086. <https://doi.org/10.1088/1742-6596/444/1/012086>.
- [27] Fredh A, Scherman JB, Fog LS, Munck af Rosenschöld P. Patient QA systems for rotational radiotherapy: a comparative experimental study with intentional errors. *Med Phys* 2013;40:031716. <https://doi.org/10.1118/1.4788645>.
- [28] 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. doi:10.1118/1.4789580.
- [29] Woon W, Ravindran PB, Ekayanake P, Vikraman S, Lim YF, Khalid J. A study on the effect of detector resolution on gamma index passing rate for VMAT and IMRT QA. *J Appl Clin Med Phys* 2018;19:230–48. <https://doi.org/10.1002/acm2.12285>.
- [30] Glenn M, Hernandez V, Saez J, Followill D, Howell R, Pollard-Larkin J, et al. Treatment plan complexity does not predict IROC Houston anthropomorphic head and neck phantom performance. *Phys Med Biol* 2018;63:205015. <https://doi.org/10.1088/1361-6560/aae29e>.
- [31] Park JM, Park S-Y, Kim H. Modulation index for VMAT considering both mechanical and dose calculation uncertainties. *Phys Med Biol* 2015;60:7101–25. <https://doi.org/10.1088/0031-9155/60/18/7101>.
- [32] Bäck A, Nordström F, Gustafsson M, Götstedt J, Hauer AK. Complexity metric based on fraction of penumbra dose – initial study. *J Phys Conf Ser* 2017;847:012002.