Physica Medica 59 (2019) 64–74

Contents lists available at ScienceDirect

Physica Medica

journal homepage: www.elsevier.com/locate/ejmp

Original paper

Evaluation of the plan delivery accuracy of intensity-modulated radiation therapy by texture analysis using fluence maps

So-Yeon Park[a,b], Jung-in Kim[b,c,d], Do Hoon Oh[e], Jong Min Park[b,c,d,f,*]

[a] Department of Radiation Oncology, Veterans Health Service Medical Center, Seoul 05368, Republic of Korea
[b] Institute of Radiation Medicine, Seoul National University Medical Research Center, Seoul 03080, Republic of Korea
[c] Biomedical Research Institute, Seoul National University Hospital, Seoul 03080, Republic of Korea
[d] Department of Radiation Oncology, Seoul National University Hospital, Seoul 03080, Republic of Korea
[e] Department of Radiation Oncology, Chung-Ang University Hospital, Seoul 06973, Republic of Korea
[f] Institute for Smart System, Robotics Research Laboratory for Extreme Environments, Advanced Institutes of Convergence Technology, Suwon 16229, Republic of Korea

ARTICLE INFO

Keywords:
Fluence
Intensity-modulated radiation therapy
Modulation index
Texture analysis

ABSTRACT

Purpose: The aim of this study is to evaluate the plan delivery accuracy of intensity-modulated radiation therapy (IMRT) by using textural features calculated from fluence maps.

Methods: In total, 202 IMRT plans were generated with the Trilogy (C-series linac) and TrueBeam STx systems. Fluence maps were obtained for each IMRT plan. With the fluence map, six textural features were calculated, namely *angular second moment (ASM)*, *inverse difference moment (IDM)*, *contrast*, *variance*, *correlation*, and *entropy* (particular displacement distances, $d = 1$, 5, and 10). To evaluate performances of textural features in predicting IMRT delivery accuracy, Spearman's rank correlation coefficients (r_s) were calculated between textural feature values and measures of conventional methods.

Results: For Trilogy, *correlation* ($d = 1$) showed r_s values of -0.764 ($p < 0.001$) with gamma passing rates of 2%/1 mm and 0.863 ($p < 0.001$) with multi-leaf collimator (MLC) errors. For TrueBeam STx, *variance* ($d = 10$) showed r_s values of 0.691 ($p < 0.001$) with gamma passing rates of 1%/2 mm and -0.585 ($p < 0.001$) with MLC errors.

Conclusion: *Correlation* ($d = 1$) for Trilogy and *variance* ($d = 10$) for TrueBeam STx demonstrated the potential in predicting IMRT delivery accuracy with high correlations to conventional IMRT delivery accuracy measures.

1. Introduction

Intensity-modulated radiation therapy (IMRT) is widely adopted in clinical settings because of its capability to deliver conformal prescription doses to target volumes and simultaneously minimize normal tissue irradiation [1–9]. To acquire this optimal dose distribution, IMRT techniques modulate photon intensities using multi-leaf collimator (MLC) movements to generate an optimal fluence map. As the level of difficulty of the desired dose distribution increases, the complexity of the fluence map also increases, which increases both the dose calculation and mechanical MLC movement uncertainties [10–14]. In this respect, highly-modulated IMRT plans may fail to deliver the desired dose distribution to a patient. Therefore, the pre-treatment verification of IMRT plan delivery accuracy, such as through patient-specific quality assurance (QA), is strongly recommended by international guidelines and is routinely performed in clinical settings [15,16].

For patient-specific QAs in general, dose distributions calculated from the treatment planning system (TPS) are compared to those measured by various measurement devices with gamma-index method. The gamma-index method suggested by Low et al. quantitatively evaluates the similarity of two dose distributions, point by point, using dose differences combined with distance-to-agreements, *i.e.,* gamma criterion [17]. The gamma passing rate is percentage of the number of measurement points which meet the gamma criterion in comparison with calculated points to the whole number of evaluated points [18]. Although the gamma-index method has been widely adopted in the clinical setting, it is limited by low sensitivity for detecting mechanical parameter errors, inaccurate dose calculations, and clinical irrelevance [19–21].

As an alternative to the gamma analysis, the modulation index (MI) was suggested by several studies as a way to verify IMRT delivery accuracy [2,11,13,18]. Compared to the measurement-based patient-specific QA methods, the modulation index has the advantage of early IMRT plan detection, through which the delivery accuracy is low

* Corresponding author at: Department of Radiation Oncology, Seoul National University Hospital, Seoul 03080, Republic of Korea.
E-mail address: leodavinci@naver.com (J.M. Park).

https://doi.org/10.1016/j.ejmp.2019.02.016
Received 3 July 2018; Received in revised form 18 February 2019; Accepted 20 February 2019
Available online 02 March 2019

Table 1
Intensity modulated radiation therapy (IMRT) Plan information.

Treatment site	*N*	Photon energy	Prescription dose (Gy)	$D_{2\%}$	$D_{98\%}$	Fraction number
Trilogy						
HN	40	6 MV	67.5, 54, 48 (for PTV1, PTV2, PTV3, respectively)	71.3 ± 1.3, 67.3 ± 1.5, 51.8 ± 2.0	66.2 ± 1.0, 52.6 ± 1.1, 47.2 ± 1.5	30
Prostate (PP)	21	15 MV	50.4	53.1 ± 0.4	48.4 ± 0.6	28
Prostate (BP)	21	15 MV	30.6	32.1 ± 0.3	29.7 ± 0.2	17
Liver	11	15 MV	50	51.9 ± 0.3	48.8 ± 0.3	20
Spine	9	15 MV	30	32.4 ± 1.4	27.4 ± 3.3	10
TrueBeam STx						
Lung SABR	20	6 MV FFF	60	68.3 ± 1.4	57.0 ± 0.5	4
Spine SABR	20	10 MV FFF	16	17.3 ± 0.6	16.0 ± 0.6	1
Liver SABR	20	10 MV FFF	42	44.7 ± 1.4	41.1 ± 0.7	3
Brain	20	6 MV	30	31.4 ± 0.7	29.4 ± 0.5	10
HN	20	6 MV	67.5, 54, 48 for PTV1, PTV2, PTV3, respectively)	71.5 ± 0.8, 67.2 ± 1.1, 52.1 ± 1.8	65.9 ± 0.4, 52.3 ± 0.7, 46.9 ± 0.7	30

Abbreviations: $D_{2\%}$ = near maximum dose, dose received by at least 2% volume of the target volume; $D_{98\%}$ = near minimum dose, dose received by at least 98% volume of the target volume; HN = head and neck; PP = primary plan; BP = boost plan; SABR = stereotactic ablative radiotherapy; FFF = flattening filter free.

because it can be calculated at the planning level. Furthermore, the modulation index can be used in the IMRT optimization process by preventing the generation of excessively modulated plans during planning. In this respect, previous studies have proposed various modulation indices. Webb et al. evaluated the complexity of IMRT plans by changes in adjacent beamlets that exceed a certain portion of a standard deviation of the fluence map [22,23]. McNiven et al. quantified variations of beam apertures and the degrees of MLC movements, introducing the modulation complexity score (MCS) [12]. Du et al. proposed modulation indices that analyse beam aperture irregularities, which include the plan averaged beam area (PA), plan averaged beam irregularity (PI), and plan averaged beam modulation (PM), for both IMRT and volumetric-modulated arc therapy (VMAT) [2]. We have also proposed modulation indices that were textural features calculated from the VMAT fluence maps [24,25]. In that study, considerable correlations were demonstrated between textural feature values and conventional patient-specific QA measures of gamma passing rates and mechanical errors during VMAT delivery, which were acquired from machine log files recorded in the linear accelerator (linac) control system.

Textural analysis on fluence maps have been suggested for evaluating the degree of VMAT modulation in previous studies [24,25]; however, this method for IMRT has not been tested. Because the principles of photon beam modulation techniques and mechanical modulation parameters of IMRT differ from those of VMAT, the feasibility of using textural features as modulation indices for IMRT should be investigated. In addition, previous studies of VMAT textural features were limited, due to only utilizing a single type of linac, a single photon energy, a single type of dosimeter, and limited treatment sites of only head and neck (HN) and prostate cancers [24,25]. More comprehensive studies that test the performance of textural features as modulation indices are required. Therefore, in this study, we evaluated the comprehensive performance of textural features that were calculated from IMRT fluence maps as modulation indices by utilizing two types of linacs, various photon energies, two diode measurement arrays with different geometries, and various treatment sites. Correlations between the values of textural features and the conventional measures of IMRT plan verification methods were investigated. The conventional measures of IMRT plan verification methods were as follows:

i) Global gamma passing rates with various gamma criteria,
ii) Differences in MLC positions between the planned and the actual delivery,
iii) Differences in dose-volumetric (DV) parameters between original IMRT plans and the plans reconstructed with actual MLC positions during plan delivery.

To acquire reliable results, we analysed a total of 202 IMRT plans in this study.

2. Materials and methods

2.1. Patient selection

A total of 181 patients with various treatment sites, including HN cancer, prostate cancer, liver cancer, lung cancer, brain tumours, and spine tumours, were retrospectively selected for this study. Approval for this study was obtained from the institutional review board (IRB No. 1802–069-922). All patients were scanned with the Brilliance CT Big Bore™ (Philips, Amsterdam, the Netherlands).

2.2. IMRT planning

For IMRT planning, two types of linacs were used: the Trilogy™ with the Millennium 120 MLC system (C-series linac), and the TrueBeam™ STx with the High-definition (HD) 120 MLC system (Varian Medical Systems, Palo Alto, CA, USA). The TrueBeam STx, which is more advanced than the C-series linacs, is an integrated system that centrally controls mechanical parameters during plan delivery using encoder signals, while the mechanical parameters of the Trilogy are controlled individually, using potentiometer signals. The most significant difference between the Trilogy and the TrueBeam STx is the MLC leaf width. The MLC leaf widths of the TrueBeam STx (2.5 mm at the central region and 5 mm at the periphery region) are half of those of the Trilogy (5 mm at the central region and 10 mm at the periphery region).

Information of IMRT plans generated in this study is summarized in Table 1. For HN IMRT plans, simultaneous integrated boost (SIB) technique was used with a total of three target volumes. The target volumes of the IMRT plans for prostate cancer, liver cancer, and spine tumor were located deep inside a body, therefore, 15 MV photon beams with a higher penetration were used for those plans.

For all IMRT plans in this study, the dose volume optimizer (DVO, ver. 13.7, Varian Medical System, Palo Alto, CA, USA) in the Eclipse™ treatment planning system (TPS) (ver. 13.7, Varian Medical System, Palo Alto, CA, USA) was used to optimize the IMRT plans. All IMRT plans in this study were optimized to meet the dose constraints stated in the Radiation Therapy Oncology Group (RTOG) protocols [26]. The monitor unit (MU) constraints and normal tissue objective (NTO) were not considered during IMRT optimization. Fluence-smoothing factors in the *x* and *y* directions were 40 and 30, respectively. After optimization, dose distributions were calculated using the anisotropic analytic algorithm (AAA, ver. 13.7, Varian Medical System, Palo Alto, CA, USA) with a calculation grid of 1 mm. All IMRT plans were normalized, such that

the dose received by 100% of the target volume was equal to or higher than 90% of the prescription dose as recommended by the International Commission on Radiation Units and Measurements (ICRU) Report 91 [27].

2.3. Calculation of textural features from fluence maps

To generate fluence maps, each IMRT plan was exported in DICOM-RT file format from the TPS. Using an in-house program written in MATLAB (R2016a, Mathworks, Inc., Natick, MA, USA), MLC positions and MUs for all segments of each field were obtained, and then the sub fluence maps for each field were generated by superposition of MUs, which were shaped by the MLCs of all segments. By summation of the sub fluence maps for an IMRT plan, fluence maps for each IMRT plan were acquired.

In order to improve the efficiency and analysis accuracy, a region of interest (ROI) in the fluence map was defined, which was an area in the fluence map with values greater than 1. With this ROI, a grey level co-occurrence matrix (GLCM), which is a matrix that represents the specific spatial relationships of the pixels of interest with specific intensity values, and its neighbour was generated, to calculate textural features. Before generating GLCMs, the fluence maps were normalized so that each beamlet in the fluence map had a grey level ranging from 0 to 127, to remove the effect of differences in the maximum MUs of IMRT plans on the GLCM sizes. Particular displacement distances (d) of 1, 5, and 10, which are offsets of a pair of pixels, were used to create GLCMs. When generating GLCMs, particular angles (θ) of 0°, 45°, 90°, and 135° were adopted. For each value of d, four GLCMs were obtained with a total of four θ angles. For each GLCM, a total of six types of textural features were calculated: *angular second moment (ASM)*, *inverse difference moment (IDM)*, *contrast*, *variance*, *correlation*, and *entropy*. The calculation methods for the six types of textural features were the same as those in previous studies [24,25]. For each value of d, the textural features of the same type, which were calculated from four GLCMs were simply averaged. Therefore, a total of 18 textural features were calculated for each IMRT plan in this study (3 values of $d \times 6$ types of textural features). The method to calculate the textural features from the fluence maps is described in detail in the supplementary material S1.

2.4. Measures of IMRT plan delivery accuracy

A total of three types of conventional verification methods for measuring IMRT delivery accuracy were adopted in this study, which included 2D global gamma evaluation, mechanical parameter analysis with machine log files recorded in the linac control system during plan delivery, and DV parameter analysis between the original IMRT plans and the IMRT plans that were reconstructed with machine log files.

For global gamma evaluation, measurements of 2D dose distributions were performed with the MapCHECK2 inserted into the MapPAHN™ and the ArcCHECK™ (Sun Nuclear Corporation, Melbourne, FL, USA). As references, the calculated dose distributions were exported from the TPS by generating verification plans with dosimeter CT images. The calculated dose distribution grid sizes were 1 mm. For both the MapCHECK2 and the ArcCHECK dosimeters, the relative sensitivity calibration of each diode and absolute dose calibration were carried out before 2D dose distribution measurements. The measured dose distributions were compared to the calculated dose distributions (references), with various gamma criteria of 3%/3 mm, 2%/2 mm, 2%/1 mm, 1%/2 mm, and 1%/1 mm. Gamma passing rates with each gamma criterion were acquired.

During IMRT plan delivery, for 2D dose distribution measurements, the actual MLC positions and corresponding MU values were recorded as machine log files in the linac control system. With the machine log files, the actual MLC positions during IMRT plan delivery were compared to the planned MLC positions, and the differences between the actual and the planned MLC positions (MLC positional errors) were calculated.

The machine log files were formatted as DICOM RT files and imported into the TPS (a pseudo-code is presented in supplementary material S2). After that, the dose distribution (assumed to be an actual dose distribution) that was given to each patient were calculated with the patient CT images and structure set, which were identical to those that were used to calculate the dose distributions of the original IMRT plan. Clinically relevant DV parameters were calculated from the IMRT plans that were reconstructed using machine log files and compared to those of the original IMRT plans. The differences between the actual and planned DV parameters were calculated.

2.5. Correlation analysis between textural features and measures of conventional IMRT plan delivery accuracy verification methods

To test the performance of textural features as modulation indices, the correlations between the textural feature values and measures of the conventional IMRT verification methods were investigated with Spearman's rank correlation coefficients (r_s) values. The corresponding p values for r_s were also calculated, to determine whether the correlation was statistically significant. When correlations were identified between textural features and DV parameter difference, numerous correlations were investigated (308 correlations were examined). Therefore, in this study, we only counted the number of DV parameters that had correlations with p values less than 0.05.

For comparisons, the conventional modulation indices that have been suggested by previous studies, which were the Webb's modulation index (MI_{Webb}), MCS, and PA were also calculated, and the correlations of these modulation indices to measures of the conventional IMRT plan verification methods were investigated for each IMRT plan.

3. Results

3.1. The values of textural features and conventional modulation indices

The values of textural features and conventional modulation indices with the Trilogy and the TrueBeam STx systems are shown in Table 2 and 3, respectively.

The HN IMRT plans are known to be highly modulated, owing to the proximity of organs at risk (OAR) to the target volumes, the concave-shaped target volumes that reduce doses to the spinal cord and brain stem, and multiple adjoining target volumes with various prescription doses. The prostate and lung SABR IMRT plans are known to be relatively lowly modulated owing to the convex or ellipsoidal target volumes, small-sized target volumes (lung SABR IMRT), and general lack of proximal OARs near the target volumes (lung SABR IMRT). Therefore, the values of *ASM*, *contrast,* and *variance* decreased as the modulation degree increased (the smallest values were observed for HN IMRT plans, while the largest values were observed for prostate or lung SABR IMRT plans). On the contrary, the values of *IDM*, *correlation*, and *entropy* increased with increases in the modulation degrees of IMRT plans (the largest values were for HN IMRT plans, while the smallest were for prostate or lung SABR plans, in general). This result was consistent with a previous study [24,25].

For conventional modulation indices, MI_{Webb} and PA should increase as the IMRT modulation degree increases, while the MCS should decrease, by their definitions [2,22,23]. The MCS and PA showed consistent tendencies with their definitions. However, for MI_{Webb}, we observed that that the spine IMRT plans were most highly modulated with the Trilogy system and the brain IMRT plans were most highly modulated with the TrueBeam STx system, which is inconsistent with the results of other modulation indices and textural features.

3.2. Correlation to the gamma passing rates

The r_s values and corresponding p values of the textural features and

Table 2
The values of textural features and conventional modulation indices for intensity modulated radiation therapy (IMRT) plans with the Trilogy system.

		HN (n = 40)	Prostate (PP) (n = 21)	Prostate (BP) (n = 21)	Liver (n = 11)	Spine (n = 9)
	d	Textural features				
ASM ($\times 10^{-3}$)	1	1.13 ± 0.24	2.24 ± 0.77	3.67 ± 1.35	2.35 ± 1.59	2.10 ± 1.11
	5	0.79 ± 0.18	2.29 ± 0.56	3.57 ± 1.16	1.65 ± 0.90	1.70 ± 0.78
	10	0.98 ± 0.26	4.33 ± 1.59	8.86 ± 5.10	2.43 ± 2.38	2.44 ± 1.36
IDM	1	0.33 ± 0.02	0.26 ± 0.02	0.28 ± 0.02	0.32 ± 0.04	0.28 ± 0.07
	5	0.15 ± 0.01	0.11 ± 0.02	0.12 ± 0.02	0.15 ± 0.03	0.13 ± 0.04
	10	0.11 ± 0.01	0.07 ± 0.01	0.08 ± 0.02	0.11 ± 0.03	0.10 ± 0.02
Contrast ($\times 10^{3}$)	1	0.10 ± 0.02	0.45 ± 0.11	0.61 ± 0.15	0.24 ± 0.15	0.33 ± 0.16
	5	0.53 ± 0.13	2.06 ± 0.51	2.00 ± 0.63	1.15 ± 0.59	1.31 ± 0.65
	10	0.83 ± 0.22	3.10 ± 0.65	3.21 ± 1.22	1.94 ± 0.84	1.97 ± 0.77
Variance	1	39.06 ± 3.54	51.74 ± 4.20	49.65 ± 5.65	49.90 ± 5.92	47.48 ± 6.91
	5	41.29 ± 3.90	55.98 ± 4.54	53.84 ± 6.04	52.62 ± 6.60	50.57 ± 7.11
	10	42.51 ± 4.22	55.47 ± 4.81	52.83 ± 8.90	53.49 ± 6.10	51.99 ± 8.11
Correlation	1	0.94 ± 0.01	0.84 ± 0.03	0.76 ± 0.05	0.91 ± 0.05	0.86 ± 0.07
	5	0.70 ± 0.06	0.38 ± 0.11	0.37 ± 0.12	0.62 ± 0.13	0.50 ± 0.25
	10	0.57 ± 0.08	0.05 ± 0.15	0.03 ± 0.17	0.37 ± 0.21	0.28 ± 0.28
Entropy	1	3.18 ± 0.06	2.84 ± 0.09	2.64 ± 0.09	2.93 ± 0.19	2.94 ± 0.15
	5	3.37 ± 0.07	2.79 ± 0.10	2.60 ± 0.10	3.02 ± 0.25	2.99 ± 0.22
	10	3.33 ± 0.07	2.57 ± 0.14	2.25 ± 0.19	2.92 ± 0.34	2.86 ± 0.29
		Conventional modulation indices				
MI_{Webb} ($\times 10^{-3}$)		4.14 ± 1.11	3.34 ± 0.71	3.18 ± 0.43	5.14 ± 2.14	5.46 ± 2.68
MCS ($\times 10^{-1}$)		0.30 ± 0.05	0.81 ± 0.14	0.94 ± 0.13	0.78 ± 0.24	0.63 ± 0.25
PA		2.27 ± 0.41	0.95 ± 0.22	0.90 ± 0.14	1.72 ± 0.60	2.01 ± 1.40

Abbreviations: HN = head and neck; PP = primary plan; BP = boost plan; *d* = particular displacement distance; ASM = angular second moment; IDM = inverse difference moment; MI_{Webb} = modulation index suggested by Webb et al.; MCS = modulation complexity score; PA = plan averaged beam area.

the conventional modulation indices to the global gamma passing rates that were obtained with the Trilogy system are shown in Table 4 and 5. Only the r_s values with p values less than 0.05 are shown in Table 4 and 5. All of the IMRT plans that were investigated in this study were clinically acceptable, with gamma passing rates higher than 95% with 3%/3 mm gamma criterion. The r_s values of the *correlation* (particular displacement distances, $d = 1$) to every gamma passing rate, regardless of the gamma criteria, were always highest for both the MapCHECK2 and ArcCHECK, *i.e.*, the correlations between the *correlation* ($d = 1$) and gamma passing rates with various gamma criteria were always strongest for both the MapCHECK2 and ArcCHECK measurements, compared to the other textural features. The strongest correlation was observed between the *correlation* ($d = 1$) and the gamma passing rate with 2%/1 mm from ArcCHECK measurements (r_s = -0.764 with $p < 0.001$). *Correlation* ($d = 10$) and PA also showed generally strong correlations to the gamma passing rates with various gamma criteria for both the MapCHECK2 and ArcCHECK measurements, although these r_s values were lower than those of *correlation* ($d = 1$). For all of the textural

Table 3
The values of textural features and conventional modulation indices for intensity modulated radiation therapy (IMRT) plans with the TrueBeam STx system. Abbreviations: SABR = stereotactic ablative radiotherapy; HN = head and neck; *d* = particular displacement distance; ASM = angular second moment; IDM = inverse difference moment; MI_{Webb} = modulation index suggested by Webb et al.; MCS = modulation complexity score; PA = plan averaged beam area.

		Lung SABR (n = 20)	Spine SABR (n = 20)	Liver SABR (n = 20)	Brain (n = 20)	HN (n = 20)
	d	Textural features				
ASM ($\times 10^{-3}$)	1	4.95 ± 1.45	1.77 ± 0.61	3.70 ± 1.38	5.92 ± 4.52	1.62 ± 0.48
	5	3.98 ± 1.54	1.60 ± 0.51	2.56 ± 1.03	3.23 ± 2.56	0.90 ± 0.30
	10	7.23 ± 3.71	2.33 ± 0.83	4.04 ± 2.23	3.53 ± 2.96	1.08 ± 0.31
IDM	1	0.29 ± 0.04	0.25 ± 0.03	0.31 ± 0.05	0.39 ± 0.07	0.41 ± 0.04
	5	0.12 ± 0.03	0.10 ± 0.01	0.13 ± 0.03	0.18 ± 0.06	0.18 ± 0.02
	10	0.11 ± 0.16	0.07 ± 0.01	0.09 ± 0.02	0.13 ± 0.05	0.13 ± 0.02
Contrast ($\times 10^{3}$)	1	0.47 ± 0.21	0.20 ± 0.09	0.29 ± 0.14	0.22 ± 0.16	0.05 ± 0.01
	5	2.50 ± 0.55	1.61 ± 0.57	1.83 ± 0.73	1.51 ± 0.93	0.40 ± 0.14
	10	3.81 ± 0.98	2.77 ± 0.96	3.05 ± 1.18	2.68 ± 1.49	0.70 ± 0.27
Variance	1	52.98 ± 4.75	46.43 ± 6.03	48.01 ± 8.17	50.65 ± 5.24	37.80 ± 3.53
	5	57.11 ± 4.59	50.08 ± 6.98	52.50 ± 9.14	55.52 ± 6.22	39.51 ± 3.76
	10	56.70 ± 10.47	51.99 ± 7.66	54.54 ± 9.48	58.47 ± 7.03	40.50 ± 4.04
Correlation	1	0.83 ± 0.08	0.91 ± 0.03	0.87 ± 0.06	0.92 ± 0.06	0.96 ± 0.01
	5	0.25 ± 0.12	0.39 ± 0.12	0.37 ± 0.14	0.55 ± 0.21	0.75 ± 0.08
	10	−0.11 ± 0.14	0.02 ± 0.15	0.03 ± 0.19	0.28 ± 0.27	0.59 ± 0.13
Entropy	1	2.54 ± 0.14	2.96 ± 0.09	2.71 ± 0.17	2.74 ± 0.25	3.07 ± 0.08
	5	2.55 ± 0.19	2.99 ± 0.13	2.78 ± 0.21	2.90 ± 0.28	3.37 ± 0.07
	10	2.30 ± 0.47	2.85 ± 0.17	2.65 ± 0.27	2.85 ± 0.30	3.37 ± 0.06
		Conventional modulation indices				
MI_{Webb} ($\times 10^{-3}$)		7.97 ± 1.41	5.62 ± 1.48	7.83 ± 2.46	9.20 ± 2.56	6.12 ± 2.24
MCS ($\times 10^{-1}$)		1.25 ± 0.18	0.64 ± 0.19	1.11 ± 0.42	1.40 ± 0.46	0.35 ± 0.08
PA		1.01 ± 0.45	1.13 ± 0.53	1.39 ± 0.67	3.72 ± 1.85	3.72 ± 0.82

Abbreviations: *d* = particular displacement distance; r_s (p) = the values of Spearman's correlation coefficient (corresponding p value); ASM = angular second moment; IDM = inverse difference moment; MI_{Webb} = modulation index suggested by Webb et al.; MCS = modulation complexity score; PA = plan averaged beam area.

Table 4
Correlations of textural features and conventional modulation indices to the global gamma passing rates from the MapCHECK2 measurements with the Trilogy system.

		3%/3 mm	2%/2 mm	2%/1 mm	1%/2 mm	1%/1 mm
	d	r_s (*p*) with the MapCHECK2 measurements				
ASM	1	0.393 (< 0.001)	0.414 (< 0.001)	0.496 (< 0.001)	0.375 (< 0.001)	0.314 (0.001)
	5	0.414 (< 0.001)	0.416 (< 0.001)	0.545 (< 0.001)	0.341 (< 0.001)	0.337 (0.001)
	10	0.414 (< 0.001)	0.418 (< 0.001)	0.555 (< 0.001)	0.347 (< 0.001)	0.364 (< 0.001)
IDM	1	−0.440 (< 0.001)	−0.367 (< 0.001)	−0.603 (< 0.001)	–	−0.395 (< 0.001)
	5	−0.421 (< 0.001)	−0.361 (< 0.001)	−0.561 (< 0.001)	–	0.361 (< 0.001)
	10	−0.438 (< 0.001)	0.393 (< 0.001)	0.567 (< 0.001)	0.226 (0.022)	0.355 (< 0.001)
Contrast	1	0.463 (< 0.001)	0.468 (< 0.001)	0.598 (< 0.001)	0.357 (< 0.001)	0.390 (< 0.001)
	5	0.382 (< 0.001)	0.372 (< 0.001)	0.554 (< 0.001)	0.252 (0.011)	0.331 (0.001)
	10	0.352 (< 0.001)	0.352 (< 0.001)	0.508 (< 0.001)	0.240 (0.015)	0.284 (0.004)
Variance	1	0.261 (0.008)	0.200 (0.044)	0.312 (0.001)	–	–
	5	0.273 (0.006)	0.226 (0.022)	0.344 (< 0.001)	–	–
	10	–	–	−0.225 (0.023)	–	–
Correlation	1	−0.520 (< 0.001)	−0.514 (< 0.001)	−0.647 (< 0.001)	−0.418 (< 0.001)	−0.469 (< 0.001)
	5	−0.413 (< 0.001)	−0.409 (< 0.001)	−0.615 (< 0.001)	−0.292 (0.003)	−0.419 (< 0.001)
	10	−0.420 (< 0.001)	−0.421 (< 0.001)	−0.627 (< 0.001)	−0.291 (0.003)	−0.413 (< 0.001)
Entropy	1	−0.451 (< 0.001)	−0.453 (< 0.001)	−0.532 (< 0.001)	−0.403 (< 0.001)	−0.413 (< 0.001)
	5	−0.443 (< 0.001)	−0.446 (< 0.001)	−0.555 (< 0.001)	−0.383 (< 0.001)	−0.341 (< 0.001)
	10	−0.452 (< 0.001)	−0.449 (< 0.001)	−0.562 (< 0.001)	−0.378 (< 0.001)	−0.361 (< 0.001)
MI_{Webb}		−0.398 (< 0.001)	−0.386 (< 0.001)	−0.347 (< 0.001)	−0.323 (0.001)	−0.229 (0.021)
MCS		0.391 (< 0.001)	0.347 (< 0.001)	0.377 (< 0.001)	0.272 (0.006)	–
PA		−0.482 (< 0.001)	−0.428 (< 0.001)	−0.645 (< 0.001)	−0.286 (0.004)	−0.437 (< 0.001)

Table 5
Correlations of textural features and conventional modulation indices to the global gamma passing rates from the ArcCHECK measurements with the Trilogy system.

		3%/3 mm	2%/2 mm	2%/1 mm	1%/2 mm	1%/1 mm
	d	r_s (*p*) with the ArcCHECK measurements				
ASM	1	0.398 (< 0.001)	0.588 (< 0.001)	0.645 (< 0.001)	0.231 (0.019)	0.474 (< 0.001)
	5	0.453 (< 0.001)	0.610 (< 0.001)	0.675 (< 0.001)	0.247 (0.012)	0.516 (< 0.001)
	10	0.482 (< 0.001)	0.633 (< 0.001)	0.707 (< 0.001)	0.279 (0.004)	0.553 (< 0.001)
IDM	1	−0.470 (< 0.001)	−0.578 (< 0.001)	−0.647 (< 0.001)	−0.301 (0.002)	−0.549 (< 0.001)
	5	−0.448 (< 0.001)	−0.529 (< 0.001)	−0.605 (< 0.001)	−0.250 (0.011)	−0.498 (< 0.001)
	10	−0.508 (< 0.001)	−0.578 (< 0.001)	0.615 (< 0.001)	−0.257 (0.009)	−0.482 (< 0.001)
Contrast	1	0.465 (< 0.001)	0.653 (< 0.001)	0.720 (< 0.001)	0.260 (0.008)	0.564 (< 0.001)
	5	0.446 (< 0.001)	0.577 (< 0.001)	0.646 (< 0.001)	–	0.482 (< 0.001)
	10	0.452 (< 0.001)	0.576 (< 0.001)	0.611 (< 0.001)	–	0.436 (< 0.001)
Variance	1	0.308 (0.002)	0.429 (< 0.001)	0.504 (< 0.001)	–	0.328 (0.001)
	5	0.323 (0.001)	0.448 (< 0.001)	0.520 (< 0.001)	–	0.341 (< 0.001)
	10	0.257 (0.009)	0.381 (< 0.001)	0.416 (< 0.001)	–	0.244 (0.013)
Correlation	1	−0.509 (< 0.001)	−0.711 (< 0.001)	−0.764 (< 0.001)	−0.386 (< 0.001)	−0.641 (< 0.001)
	5	−0.492 (< 0.001)	−0.607 (< 0.001)	−0.665 (< 0.001)	−0.264 (0.007)	−0.540 (< 0.001)
	10	−0.500 (< 0.001)	−0.608 (< 0.001)	−0.661 (< 0.001)	−0.219 (0.027)	−0.508 (< 0.001)
Entropy	1	−0.425 (< 0.001)	−0.615 (< 0.001)	−0.683 (< 0.001)	−0.231 (0.019)	−0.499 (< 0.001)
	5	−0.458 (< 0.001)	−0.634 (< 0.001)	−0.706 (< 0.001)	−0.248 (0.012)	−0.532 (< 0.001)
	10	−0.476 (< 0.001)	−0.649 (< 0.001)	−0.722 (< 0.001)	−0.269 (0.006)	−0.556 (< 0.001)
MI_{Webb}		−0.271 (0.006)	−0.346 (< 0.001)	−0.355 (< 0.001)	–	−0.321 (0.001)
MCS		0.426 (< 0.001)	0.589 (< 0.001)	0.660 (< 0.001)	–	0.450 (< 0.001)
PA		−0.505 (< 0.001)	−0.667 (< 0.001)	−0.739 (< 0.001)	−0.370 (< 0.001)	−0.627 (< 0.001)

Abbreviations: *d* = particular displacement distance; r_s (*p*) = the values of Spearman's correlation coefficient (corresponding *p* value); ASM = angular second moment; IDM = inverse difference moment; MI_{Webb} = modulation index suggested by Webb et al.; MCS = modulation complexity score; PA = plan averaged beam area.

features, the gamma passing rates with 2%/1 mm generally showed stronger correlations than the other gamma passing rates. Except for the gamma passing rates with 1%/2 mm, every gamma passing rate showed statistically significant correlations to every textural feature, with *p* values less than 0.05.

The r_s values and corresponding *p* values of the textural features and the conventional modulation indices of the global gamma passing rates that were obtained with the TrueBeam STx are shown in Table 6 and 7. The r_s values with *p* values of < 0.05 are shown in Table 6 and 7. No textural feature or conventional modulation index consistently showed the strongest correlation to every gamma passing rate, with various gamma criteria. In addition, the correlations of the MapCHECK2 measurements differed from those of the ArcCHECK2 measurements. For the MapCHECK2 measurements, the MCS showed the strongest correlations to the gamma passing rates with 3%/3 mm and 2%/2 mm (relatively loose gamma criteria), while *variance* ($d = 10$) showed the strongest correlations to the gamma passing rates with 1%/2 mm and 1%/1 mm (relatively tight gamma criteria). For the ArcCHECK measurements, the MCS showed the strongest correlations with every gamma passing rate, except the gamma passing rates with 2%/1 mm. The *variance* ($d = 10$) showed the second strongest correlations with every gamma passing rate, except the gamma passing rates with 2%/1 mm. The strongest correlation was observed between the MCS and the gamma passing rates with 1%/2 mm ($r_s = 0.699$ with $p < 0.001$). Statistically significant correlations were more frequently observed with the MapCHECK2 measurements than with the ArcCHECK measurements.

Table 6
Correlations of textural features and conventional modulation indices to the global gamma passing rates from the MapCHECK2 measurements with the TrueBeam STx system.

		3%/3 mm	2%/2 mm	2%/1 mm	1%/2 mm	1%/1 mm
	d	r_s (p) with the MapCHECK2 measurements				
ASM	1	0.232 (0.020)	0.276 (0.005)	0.206 (0.040)	–	–
	5	0.345 (< 0.001)	0.379 (< 0.001)	0.293 (0.003)	0.273 (0.006)	0.264 (0.008)
	10	0.370 (< 0.001)	0.384 (< 0.001)	0.299 (0.002)	0.277 (0.005)	0.258 (0.009)
IDM	1	−0.227 (0.023)	−0.275 (0.006)	−0.235 (0.019)	−0.256 (0.010)	−0.222 (0.026)
	5	−0.204 (0.041)	−0.276 (0.005)	−0.256 (0.010)	−0.259 (0.009)	−0.240 (0.016)
	10	−0.216 (0.031)	−0.265 (0.008)	−0.233 (0.020)	−0.328 (0.001)	−0.300 (0.002)
Contrast	1	0.356 (< 0.001)	0.432 (< 0.001)	0.357 (< 0.001)	0.350 (< 0.001)	0.336 (0.001)
	5	0.405 (< 0.001)	0.503 (< 0.001)	0.451 (< 0.001)	0.419 (< 0.001)	0.420 (< 0.001)
	10	0.359 (< 0.001)	0.458 (< 0.001)	0.408 (< 0.001)	0.488 (< 0.001)	0.477 (< 0.001)
Variance	1	0.403 (< 0.001)	0.515 (< 0.001)	0.491 (< 0.001)	0.403 (0.001)	0.438 (< 0.001)
	5	0.409 (< 0.001)	0.532 (< 0.001)	0.519 (< 0.001)	0.449 (< 0.001)	0.486 (< 0.001)
	10	0.345 (< 0.001)	0.465 (< 0.001)	0.451 (< 0.001)	0.522 (< 0.001)	0.540 (< 0.001)
Correlation	1	−0.293 (0.003)	−0.313 (0.002)	−0.235 (0.018)	−0.259 (0.009)	−0.231 (0.021)
	5	−0.293 (0.003)	−0.331 (0.001)	−0.257 (0.010)	−0.285 (0.004)	−0.252 (0.012)
	10	−0.280 (0.005)	−0.308 (0.002)	−0.243 (0.015)	−0.318 (0.001)	−0.284 (0.004)
Entropy	1	−0.335 (< 0.001)	−0.392 (< 0.001)	−0.312 (0.002)	−0.260 (0.009)	−0.260 (0.009)
	5	−0.383 (< 0.001)	−0.435 (< 0.001)	−0.344 (< 0.001)	−0.323 (0.001)	−0.305 (0.002)
	10	−0.391 (< 0.001)	−0.433 (< 0.001)	−0.346 (< 0.001)	−0.330 (0.001)	−0.307 (0.002)
MI_{Webb}		−0.241 (0.016)	−0.200 (0.046)	–	–	–
MCS		0.453 (< 0.001)	−0.548 (< 0.001)	0.509 (< 0.001)	0.460 (< 0.001)	0.486 (< 0.001)
PA		−0.223 (0.026)	−0.236 (0.018)	–	–	–

Abbreviations: d = particular displacement distance; r_s (p) = the values of Spearman's correlation coefficient (corresponding p value); ASM = angular second moment; IDM = inverse difference moment; MI_{Webb} = modulation index suggested by Webb et al.; MCS = modulation complexity score; PA = plan averaged beam area.

3.3. Correlations of the MLC positional errors

Correlations of textural features and conventional modulation indices to the MLC positional errors are shown in Table 8. For the Trilogy system, the mean MLC error was 0.25 mm ± 0.04 mm, while the mean MLC error of the TrueBeam STx system was 0.11 mm ± 0.06 mm. For both the Trilogy and TrueBeam STx systems, every correlation between the textural features and the MLC errors was statistically significant, with p values of < 0.05, except for that of *ASM* (d = 1) with the TrueBeam STx MLC errors. For the Trilogy system, the r_s values of *ASM* (d = 1, 5, and 10), *contrast* (d = 1 and 5), *correlation* (d = 1, 5, and 10), *entropy* (d = 1, 5, and 10), and PA were higher than 0.8 (all with $p < 0.001$), showing strong correlations with MLC errors. In the case of the TrueBeam STx system, the r_s values were lower than those of the Trilogy system. The *variance* (d = 1, 5, and 10) and PA showed r_s values higher than 0.55 (all with $p < 0.001$). The PA showed the strongest correlations with MLC errors, among all factors (r_s = 0.909 with $p < 0.001$ for the Trilogy system and r_s = 0.841 with $p < 0.001$ for the TrueBeam STx system).

Table 7
Correlations of textural features and conventional modulation indices to the global gamma passing rates from the ArcCHECK measurements with the TrueBeam STx system.

		3%/3 mm	2%/2 mm	2%/1 mm	1%/2 mm	1%/1 mm
	d	r_s (p) with the ArcCHECK measurements				
ASM	1	0.223 (0.026)	0.336 (0.001)	0.290	0.454 (< 0.001)	0.473 (< 0.001)
	5	–	0.289 (0.004)	–	0.427 (< 0.001)	0.449 (< 0.001)
	10	–	0.249 (0.012)	–	0.389 (< 0.001)	0.406 (< 0.001)
IDM	1	−0.236 (0.018)	–	−0.249 (0.012)	–	–
	5	–	–	−0.208 (0.038)	–	–
	10	–	–	–	–	–
Contrast	1	–	0.281 (0.005)	–	0.416 (< 0.001)	0.438 (< 0.001)
	5	–	0.375 (< 0.001)	–	0.501 (< 0.001)	0.470 (< 0.001)
	10	–	0.418 (< 0.001)	–	0.539 (< 0.001)	0.477 (< 0.001)
Variance	1	0.336 (0.001)	0.508 (< 0.001)	0.302 (0.002)	0.617 (< 0.001)	0.545 (< 0.001)
	5	0.353 (< 0.001)	0.553 (< 0.001)	0.322 (0.001)	0.657 (< 0.001)	0.579 (< 0.001)
	10	0.422 (< 0.001)	0.601 (< 0.001)	0.396 (< 0.001)	0.691 (< 0.001)	0.611 (< 0.001)
Correlation	1	–	–	–	−0.244 (0.014)	−0.295 (0.003)
	5	–	–	–	−0.248 (0.013)	−0.262 (0.008)
	10	–	–	–	−0.250 (0.012)	−0.242 (0.015)
Entropy	1	–	−0.345 (< 0.001)	−0.242 (0.015)	−0.479 (< 0.001)	−0.500 (< 0.001)
	5	–	−0.319 (0.001)	–	−0.465 (< 0.001)	−0.474 (< 0.001)
	10	–	−0.283 (0.004)	–	−0.425 (< 0.001)	−0.430 (< 0.001)
MI_{Webb}		−0.375 (< 0.001)	−0.341 (0.001)	−0.518 (< 0.001)	−0.292 (0.003)	−0.380 (< 0.001)
MCS		0.431 (< 0.001)	0.602 (< 0.001)	0.451 (< 0.001)	0.699 (< 0.001)	0.672 (< 0.001)
PA		−0.235 (0.019)	–	−0.225 (0.024)	–	–

Abbreviations: d = particular displacement distance; r_s (p) = the values of Spearman's correlation coefficient (corresponding p value); ASM = angular second moment; IDM = inverse difference moment; MI_{Webb} = modulation index suggested by Webb et al.; MCS = modulation complexity score; PA = plan averaged beam area.

Table 8
Correlations of textural features and conventional modulation indices to the differences in multi-leaf collimator (MLC) positions with the Trilogy and TrueBeam STx systems.

		Trilogy		TrueBeam STx	
	d	r_s	*p*	r_s	*p*
ASM	1	−0.814	< 0.001	–	–
	5	−0.875	< 0.001	−0.324	0.001
	10	−0.873	< 0.001	−0.425	< 0.001
IDM	1	0.719	< 0.001	0.468	< 0.001
	5	0.668	< 0.001	0.432	< 0.001
	10	0.672	< 0.001	0.370	< 0.001
Contrast	1	−0.859	< 0.001	−0.362	< 0.001
	5	−0.805	< 0.001	−0.336	0.001
	10	−0.749	< 0.001	−0.232	0.020
Variance	1	−0.617	< 0.001	−0.564	0.001
	5	−0.641	< 0.001	−0.576	< 0.001
	10	−0.499	< 0.001	−0.585	< 0.001
Correlation	1	0.863	< 0.001	0.398	< 0.001
	5	0.808	< 0.001	0.479	< 0.001
	10	0.826	< 0.001	0.477	< 0.001
Entropy	1	0.858	< 0.001	0.281	< 0.001
	5	0.892	< 0.001	0.365	< 0.001
	10	0.891	< 0.001	0.419	< 0.001
MI_{Webb}		0.213	0.032	0.411	< 0.001
MCS		−0.727	< 0.001	–	–
PA		0.909	< 0.001	0.841	< 0.001

Abbreviations: *d* = particular displacement distance; r_s = the values of Spearman's correlation coefficient; ASM = angular second moment; IDM = inverse difference moment; MI_{Webb} = modulation index suggested by Webb et al.; MCS = modulation complexity score; PA = plan averaged beam area.

3.4. Correlations of the differences in dose-volumetric parameters between original IMRT plans and the IMRT plans reconstructed with machine log files

The percent of DV parameters that showed statistically significant correlations ($p < 0.05$) between the differences in the actually delivered DV parameters compared to the planned DV parameters and textural features (and conventional modulation indices) to the whole DV parameters tested in this study are shown in Fig. 1. A total of 152 and 156 DV parameters from IMRT plans with the Trilogy and TrueBeam STx systems were investigated, respectively. For the Trilogy, *correlation* ($d = 10$) showed correlations ($p < 0.05$) to the DV parameter errors most frequently (31.6%). Among the conventional modulation indices, MCS showed correlations ($p < 0.05$) to the DV parameter errors most frequently (21.1%). For the TrueBeam STx system, *ASM* ($d = 10$) and MCS showed correlations ($p < 0.05$) to the DV parameter errors most frequently (both 16.0%).

4. Discussion

The values of textural features calculated from IMRT plans with various treatment sites in the present study were in close agreement with those found in previous studies that analysed VMAT plans [24,25]. The values of *ASM*, *contrast*, and *variance* were small for the highly-modulated IMRT plans (HN plans) and large for the lowly-modulated IMRT plans (prostate or lung SABR plans). On the contrary, the values of *IDM*, *correlation*, and *entropy* showed large values for the highly-modulated IMRT plans and small values for the lowly-modulated IMRT plans. This tendency was similar to those observed in previous studies of VMAT plans [24,25]. Therefore, the changes in the textural feature values according to the modulation degree of IMRT were similar to those of VMAT. For the conventional modulation indices, the changes in the values according to the modulation degree of IMRT in the present study were also similar to those of previous studies, which showed that MI_{webb} and PA increased, while MCS decreased, as the IMRT modulation degree increased [2,13,22]. Although the changes of textural features were similar to those of previous studies, the performances of each textural feature differed from those of previous studies [24,25].

In the previous study, *contrast* ($d = 1$) and *variance* ($d = 1$) showed the best performances for predicting VMAT delivery accuracy [24,25]. On the contrary, in this study, *correlation* ($d = 1$) showed the best performance, among all the tested textural features, in predicting gamma passing rates. *Correlation* ($d = 10$) showed the best performance towards predicting DV parameter errors with the Trilogy system. Both the *correlation* ($d = 1$) and *correlation* ($d = 10$) showed generally good performance in predicting MLC errors of the Trilogy system, with r_s values higher than 0.82 ($p < 0.001$). For the TrueBeam STx system, *variance* ($d = 10$) generally showed the best performance, with generally strong correlations to various conventional verification methods. Although the previous study was performed with a C-series linac and MapCHECK2 dosimeter, the *correlation* ($d = 1$) showed the best

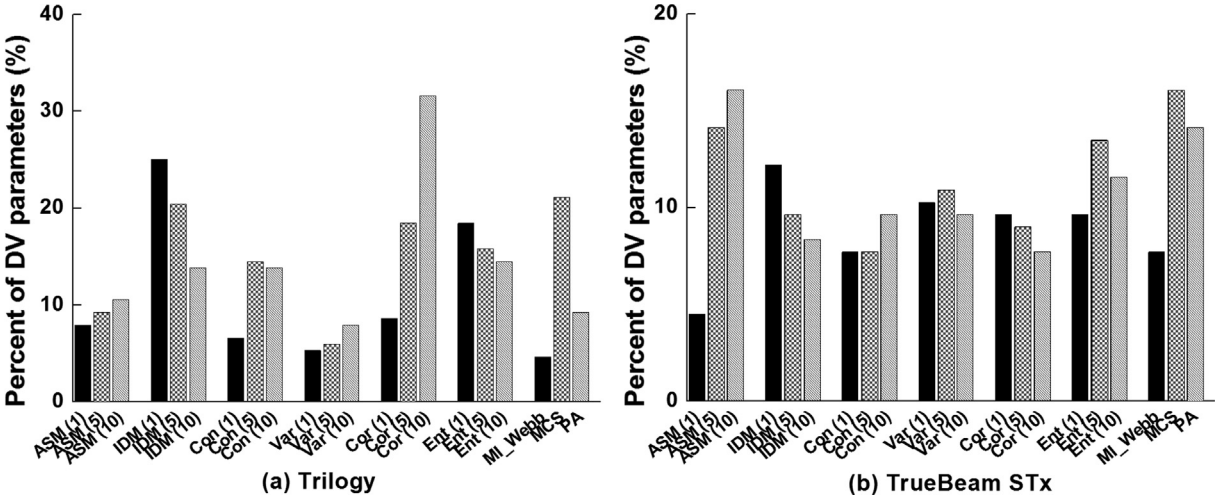

Fig. 1. The percent of dose-volumetric (DV) parameters with statistically significant correlations ($p < 0.05$) of the textural features and conventional modulation indices to the differences in DV parameters, between original intensity modulated radiation therapy (IMRT) plans and the IMRT plans reconstructed with machine log files. The textural features were *angular second moment* (*ASM*), *inverse difference moment* (*IDM*), *contrast* (*Con*), *variance* (*Var*), *correlation* (*Cor*), and *entropy* (*Ent*), with particular displacement distances of 1, 5, and 10. The results of the conventional modulation indices, which were the modulation index suggested by Webb et al. (MI_Webb), the modulation complexity score (MCS), and the plan-averaged beam area (PA), are also shown. A total of 152 and 156 DV parameters from IMRT plans with the Trilogy and TrueBeam STx systems were investigated, respectively.

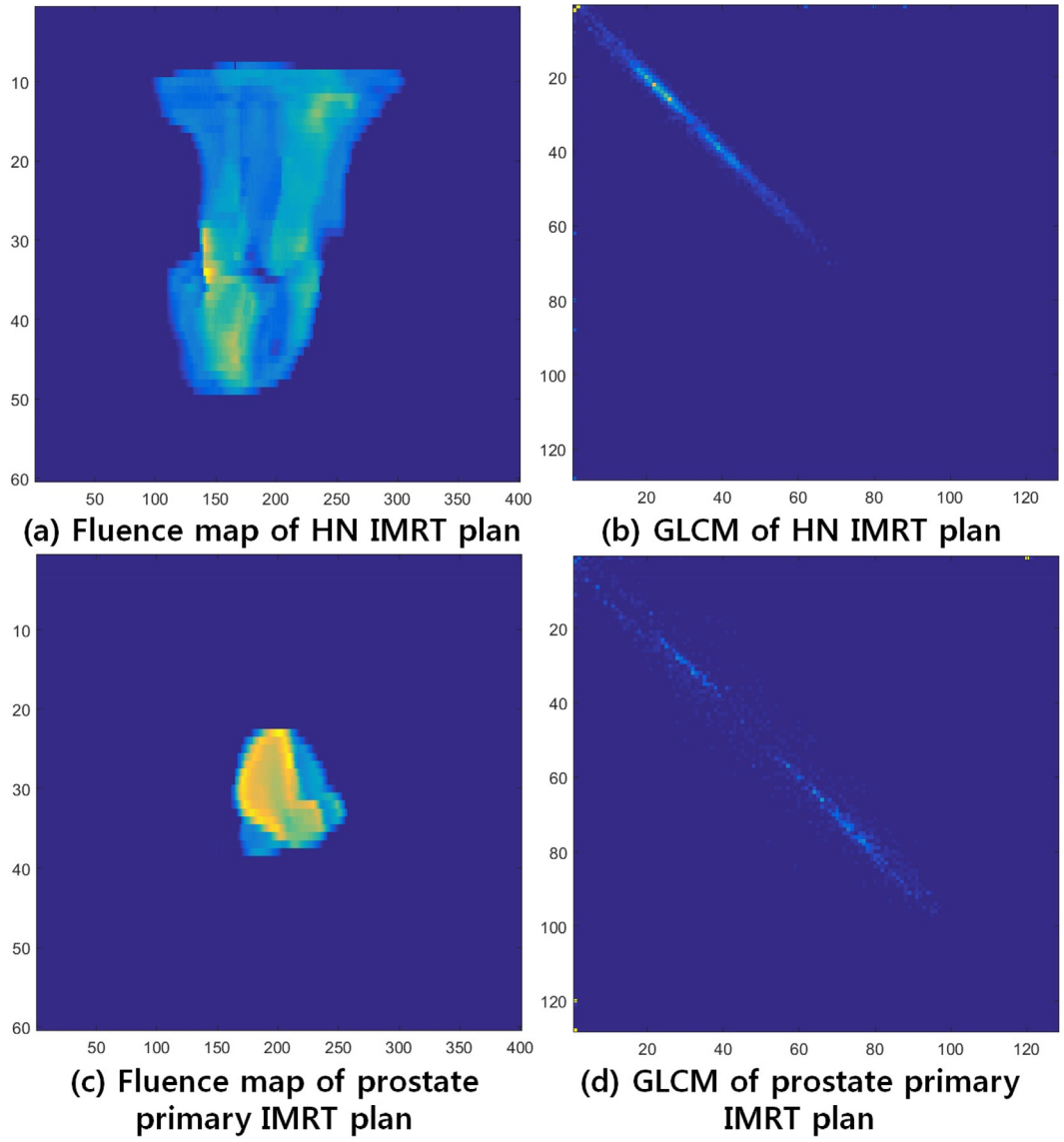

Fig. 2. The fluence map of head and neck (HN) (a) and prostate primary (c) intensity modulated radiation therapy (IMRT) plans with the Trilogy system are shown. The grey level co-occurrence matrices (GLCMs) of HN (b) and prostate primary (d) IMRT plans with the Trilogy system are also shown.

performance with the C-series linac (Trilogy) and the MapCHECK2 dosimeter in this study. Therefore, the capability of textural features to predict IMRT delivery accuracy differs from VMAT delivery accuracy predictions, owing to different delivery techniques between IMRT and VMAT.

The textural feature that showed the strongest correlations to the measures of the conventional IMRT verification methods with the Trilogy system differed from those of the TrueBeam STx system (*correlation* with $d = 1$ and 10 for the Trilogy vs. *variance* with $d = 10$ for the TrueBeam STx). Because the MLC leaf widths of the Trilogy system differ from those of the TrueBeam STx system, the resolution of the fluence maps of the Trilogy system differ from those of the TrueBeam STx system, *i.e.,* higher resolution of the fluence maps of the TrueBeam STx system than those of the Trilogy system. This difference could affect the textural feature values and impact the capacity of textural features to predict IMRT delivery accuracy. According to the results of this study, we recommend *correlation* ($d = 1$ or 10) for use as a modulation index for the C-series linac with the Millennium 120 MLC and *variance* ($d = 10$) as a modulation index for the TrueBeam STx with the HD 120 MLC.

While reviewing the MLC errors, we found that correlations of the textural features to the MLC errors of the TrueBeam STx system were always weaker than those of the Trilogy system. In this study, the mean MLC error of the TrueBeam STx system was much smaller than that of the Trilogy system (0.11 mm vs. 0.25 mm), because the TrueBeam STx beam delivery system is more accurate and precise than that of the Trilogy (C-series linac). Therefore, the ranges of MLC errors of the TrueBeam STx system were smaller than those of the Trilogy system. This could explain the weak correlations of textural features with the MLC errors of the TrueBeam STx system that were observed in this study. In an ideal radiotherapy machine with no MLC errors, there might be no correlations between the textural features and MLC errors. This can also be a reason why DV parameter errors showed statistically significant correlations with the TrueBeam STx textural features, which were generally smaller than those of the Trilogy system. Actually, the magnitudes of the differences in the DV parameters between the original IMRT plans and the IMRT plans reconstructed with the machine log files were generally smaller with the TrueBeam STx system than with the Trilogy system (data are not shown). The accurate and precise IMRT delivery capability of the TrueBeam STx system might result in generally weaker correlations of the textural features to the measures of the IMRT verification methods, compared to the results of the Trilogy

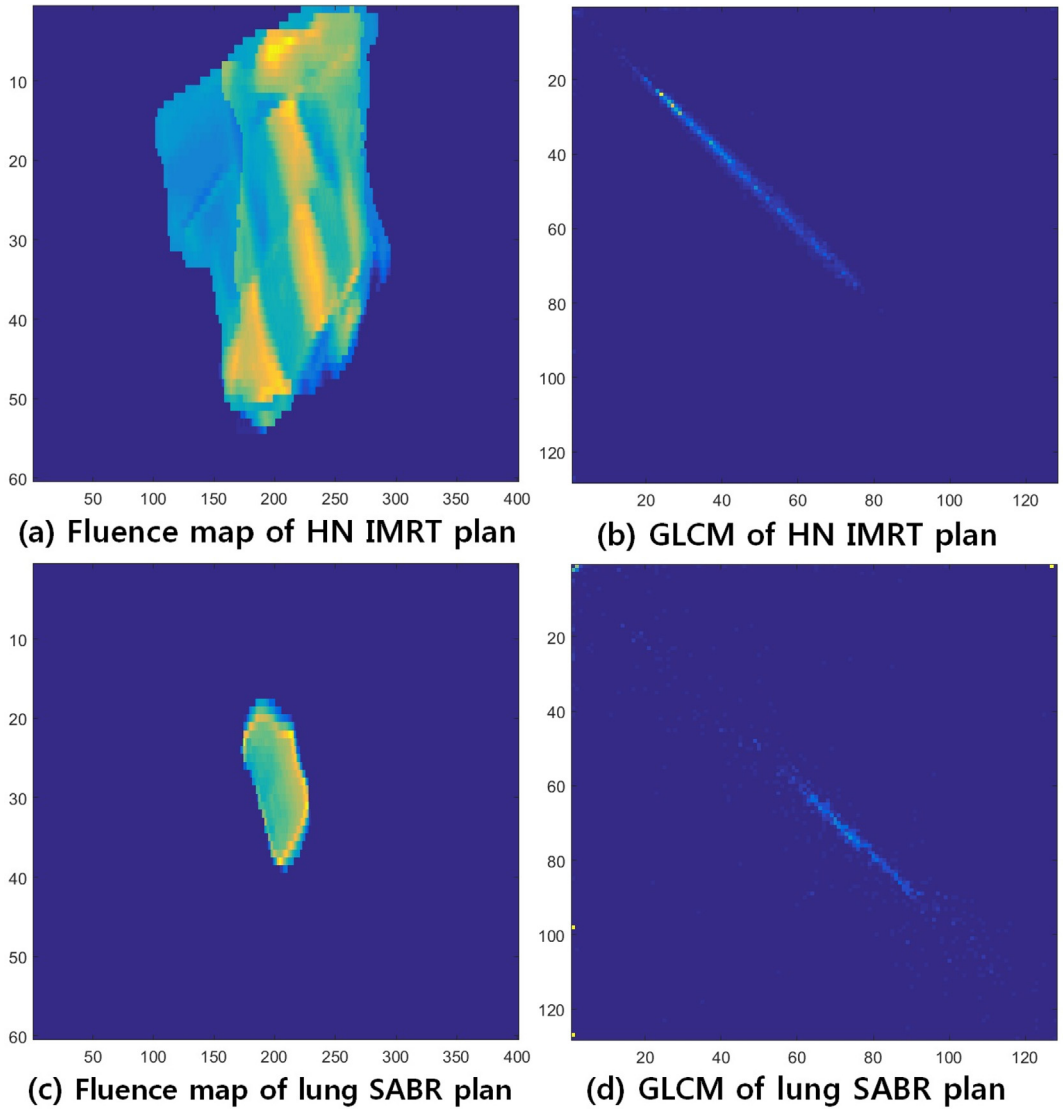

Fig. 3. The fluence map of head and neck (HN) intensity modulated radiation therapy (IMRT) (a) and lung stereotactic ablative radiotherapy (SABR) (c) IMRT plans with the TrueBeam STx system are shown. The grey level co-occurrence matrices (GLCMs) of HN IMRT (b) and lung SABR (d) IMRT plans with the TrueBeam STx system are also shown.

system.

Compared to results of the previous studies [24,25], the values of r_s between the textural features and gamma passing rates with various gamma criteria were relatively low in this study. This might be caused by the various treatment sites and various energies of this study. In addition, the number of IMRT plan segments was generally four times greater than that of VMAT plans. Owing to the fluence map generated by a huge number of segments, the modulation degree information could be smeared out for IMRT compared to VMAT. Nevertheless, some textural features showed stronger correlations than the conventional modulation indices to gamma passing rates (*correlation* with $d = 1$ and *variance* with $d = 10$), having the possibility of evaluating the plan deliverability as a modulation index. In the case of MLC errors, although the PA exhibited the highest correlations to MLC errors, *correlation* ($d = 1$) and *variance* ($d = 10$) also showed strong correlations to MLC errors. In addition, they showed stronger correlations to the MLC errors than did the MCS and MI_{webb}. Because the gamma-index method and machine log file analysis have their own limitations, and there is no golden reference to accurately represent IMRT delivery accuracy, no single indicator will always show the strongest correlations to every measure of the conventional IMRT verification methods. To review the results comprehensively, *correlation* ($d = 1$ and 10) and *variance* ($d = 10$) can be recommended to be used as modulation indices.

In this study, the gamma-index method was used as one of the verification methods to evaluate IMRT plan delivery accuracy. As several studies reported that the sensitivity of the gamma-index method was variable to detect IMRT delivery errors and the gamma passing rates depended on the type of dosimeters, measurement methods (true composite, perpendicular field-by-field, and perpendicular composite), and dose calculation algorithms [28–31]. In this regard, the American Association of Physicists in Medicine (AAPM) task group (TG) 218 report recommends that the gamma passing rates should not be used alone but the comprehensive evaluation utilizing multiple gamma criteria, gamma distributions, and gamma values together should be performed when using the gamma-index method [31]. Therefore, we might not observe consistent strong correlations between the textural feature values and the gamma passing rates alone in the present study. In the future, further study on the textural features in relation to the comprehensive results of the gamma-index method following the AAPM TG 218 will be performed.

The fluence maps and grey level co-occurrence matrices (GLCMs) of the HN IMRT, prostate primary IMRT, and lung SABR IMRT plans are

shown in Figs. 2 and 3. The fluence maps of HN IMRT plans (highly-modulated IMRT plans) had small variations between neighbouring beamlet intensities, owing to the summation of numerous irregular and small fields over a wide area. The fluence maps of prostate primary IMRT and lung SABR IMRT plans (lowly-modulated IMRT plans) showed good homogeneities between beamlets with high intensities inside the a ROI areas of fluence maps. They also showed steep gradients at the border regions. In this respect, GLCMs of HN IMRT plans showed more linearity and smaller values than those of prostate primary IMRT and lung SABR IMRT plans. Therefore, the *ASM* values of prostate primary IMRT and lung SABR IMRT plans were higher than those of HN IMRT plans. On the other hand, the fluence maps of prostate primary IMRT and lung SABR IMRT plans were more uniform. Because the *IDM* and *contrast* measure the contributions of pixels away from the diagonal in the GLCM (proportionally for the *IDM* while inversely for the *contrast*), the *IDM* values of HN IMRT plans were high, while those of *contrast* were low. *Variance* puts relatively high weights on the pixels that differ from the mean value of GLCMs. Therefore, the values of *variance* of HN IMRT plans were lower than those of prostate primary IMRT and lung SABR IMRT plans. Because *correlation* measures how correlated a pixel is to its neighbours over a whole fluence map, which ranges from 1 to −1, the mean values of *correlation* ($d = 10$) for lung SABR IMRT plans had negative signs. *Entropy* is a measure of the disorder and randomness of a fluence map. For example, white noise images result in high entropy values. Therefore, HN IMRT plans that have inhomogeneous regions inside ROIs of the fluence maps show higher values than the prostate primary IMRT and lung SABR IMRT plans.

Mathematical procedures to characterize textures fall into two major categories: statistical and syntactic methods. In this study, we used statistical methods to analyse the spatial distribution of grey values by using GLCM and then calculated a total of six textural features. Besides the GLCM-based methods, which are popular in the field of texture analysis, there are many other statistical feature-extraction methods, such as the grey level run length matrix (GLRLM) [32], the grey level difference matrix (GLDM) [33,34], and the grey level size zone matrix (GLSZM) [35]. Hundreds or more textural features can be derived from these matrices. These matrices can capture higher order statistical relations between grey values in pixel pairs or groups of pixels, in order to estimate their probability density functions [36]. Their validity in the medical field has been proved in several studies. Mir et al. detected abnormalities in CT images using GLRLM and GLDM [37]. Kim et al. utilized GLRLM and GLDM to detect micro-calcifications in digitized mammograms [38]. More optimal textual features can be used for modulation indices when hundreds of textural features are analysed, and this will be performed in the future. By identifying an optimal textural feature, or modifying and combining some textural features, to generate an optimal indicator for predicting IMRT delivery accuracy, the generation of IMRT plans with high delivery accuracy could be enabled by inserting these textural features into the planning process. This will be investigated in the future.

The textural features could aid to evaluate IMRT delivery accuracy, however, we cannot claim that the textural features can completely replace the conventional patient-specific QA based on measurements since they do not verify the final delivery accuracy of IMRT plans. For example, calculation of the textural features cannot detect network errors in the record and verify system or malfunction of the linac. The textural features can be used as an auxiliary tool at planning level to detect undeliverable IMRT plans owing to excessive modulations.

5. Conclusions

The *correlation* ($d = 1$ and 10) for the C-series linac and *variance* ($d = 10$) for the TrueBeam STx system showed excellent predictions of IMRT plan delivery accuracies. These textural features could be used as auxiliary tools to perform patient-specific QA and generate IMRT plans with high delivery accuracy at the planning level. The proposed verification method for IMRT delivery accuracy in this study is beneficial for reducing resources in the clinical setting, because it can be simply calculated with IMRT plans.

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. 2017M2A2A7A02020639, No. 2017M2A2A7A02020640, No. 2017M2A2A7A02020641, and No. 2017M2A2A7A02020643).

Conflicts of interest

The authors have no competing interests to declare.

References

[1] Luxton G, Hancock SL, Boyer AL. Dosimetry and radiobiologic model comparison of IMRT and 3D conformal radiotherapy in treatment of carcinoma of the prostate. Int J Radiat Oncol Biol Phys 2004;59:267–84.
[2] Hong L, Hunt M, Chui C, Spirou S, Forster K, Lee H, et al. Intensity-modulated tangential beam irradiation of the intact breast. Int J Radiat Oncol Biol Phys 1999;44:1155–64.
[3] Young R, Snyder B. IMRT (intensity modulated radiation therapy): progress in technology and reimbursement. Radiol Manage 2001;23:20–6.
[4] Kim SY. Reirradiation of head and neck cancer in the era of intensity-modulated radiotherapy: patient selection, practical aspects, and current evidence. Radiat Oncol J. 2017;35:1–15.
[5] Mani KR, Upadhayay S, Maria Das KJ. Influence of jaw tracking in intensity-modulated and volumetric-modulated arc radiotherapy for head and neck cancers: a dosimetric study. Radiat Oncol J 2017;35:90–100.
[6] Sakanaka K, Itasaka S, Ishida Y, Fujii K, Horimatsu T, Mizowaki T, et al. Dosimetric advantages and clinical outcomes of simultaneous integrated boost intensity-modulated radiotherapy for anal squamous cell carcinoma. Radiat Oncol J 2017;35:368–79.
[7] Cho B. Intensity-modulated radiation therapy: a review with a physics perspective. Radiat Oncol J 2018;36:1–10.
[8] Choi KH, Kim J, Lee SW, Kang YN, Jang H. Dosimetric comparison between modulated arc therapy and static intensity modulated radiotherapy in thoracic esophageal cancer: a single institutional experience. Radiat Oncol J 2018;36:63–70.
[9] Rastogi K, Sharma S, Gupta S, Agarwal N, Bhaskar S, Jain S. Dosimetric comparison of IMRT versus 3DCRT for post-mastectomy chest wall irradiation. Radiat Oncol J 2018;36:71–8.
[10] Park JM, Park SY, Kim H, Kim JH, Carlson J, Ye SJ. Modulation indices for volumetric modulated arc therapy. Phys Med Biol 2014;59:7315–40.
[11] Park JM, Park SY, Kim H. Modulation index for VMAT considering both mechanical and dose calculation uncertainties. Phys Med Biol 2015;60:7101–25.
[12] McNiven AL, Sharpe MB, Purdie TG. A new metric for assessing IMRT modulation complexity and plan deliverability. Med Phys 2010;37:505–15.
[13] Du W, Cho SH, Zhang X, Hoffman KE, Kudchadker RJ. Quantification of beam complexity in intensity-modulated radiation therapy treatment plans. Med Phys 2014;41:021716.
[14] Song J. Dosimetric analysis on the effect of target motion in the delivery of conventional IMRT, RapidArc and Tomotherapy. Prog Med Phys 2017;28:164–70.
[15] Song J. Evaluation of dynamic delivery quality assurance process for internal target volume based RapidArc. Prog Med Phys 2017;28:181–9.
[16] Oh Y, Shin DO, Kim J, Kwon J, Lee SS, Choi SH, et al. Proposal on guideline for quality assurance of radiation treatment planning system. Prog Med Phys 2017;28:191–206.
[17] Van Dyk J, Barnett RB, Cygler JE, Shragge PC. Commissioning and quality assurance of treatment planning computers. Int J Radiat Oncol Biol Phys 1993;26:261–73.
[18] Park S, Park JM, Choi CH, Chun M, Han JH, Cho JD, et al. Optimal density assignment to 2D diode array detector for different dose calculation algorithms in patient specific VMAT QA. J Radiat Prot Res 2017;42:9–15.
[19] Tyagi N, Yang K, Yan D. Comparing measurement-derived (3DVH) and machine log file-derived dose reconstruction methods for VMAT QA in patient geometries. J Appl Clin Med Phys 2014;15:4645.
[20] Kim JI, Park SY, Kim HJ, Kim JH, Ye SJ, Park JM. The sensitivity of gamma-index method to the positioning errors of high-definition MLC in patient-specific VMAT QA for SBRT. Radiat Oncol 2014;9:167.
[21] 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.
[22] Webb S. Optimization by simulated annealing of three-dimensional conformal treatment planning for radiation fields defined by a multileaf collimator. Phys Med Biol 1992;36:1201–26.
[23] Webb S. Optimization by simulated annealing of three-dimensional, conformal treatment planning for radiation fields defined by a multileaf collimator: II.

Inclusion of two-dimensional modulation of the x-ray intensity. Phys Med Biol 1992;37:1689–704.
[24] Park SY, Kim IH, Ye SJ, Carlson J, Park JM. Texture analysis on the fluence map to evaluate the degree of modulation for volumetric modulated arc therapy. Med Phys 2014;41:111718.
[25] Park SY, Park JM, Sung W, Kim IH, Ye SJ. Texture analysis on the edge-enhanced fluence of VMAT. Radiat Oncol 2015;10:74.
[26] Radiation Oncology/Toxicity/RTOG, https://en.wikibooks.org/wiki/Radiation_Oncology/Toxicity/RTOG [accessed 15 July 2009].
[27] ICRU Report 91: Prescribing, Recording, and Reporting of Stereotactic Treatments with Small Photon Beams. J ICRU 2017;14.
[28] Eaton DJ, Tyler J, Backshall A, Bernstein D, Carver A, Gasnier A, et al. An external dosimetry audit programme to credential static and rotational IMRT delivery for clinical trials quality assurance. Phys Med 2017;35:25–30.
[29] Deshpande S, Geurts M, Vial P, Metcalfe P, Lee M, Holloway L. Clinical significance of treatment delivery errors for helical TomoTherapy nasopharyngeal plans – a dosimetric simulation study. Phys Med 2017;33:159–69.
[30] Pogson EM, Aruguman S, Hansen CR, Currie M, Oborn BM, Blake SJ, et al. Multi-institutional comparison of simulated treatment delivery errors in ssIMRT, manually planned VMAT and autoplan-VMAT plans for nasopharyngeal radiotherapy. Phys Med 2017;42:55–6.
[31] Miften M, Olch A, Mihailidis D, Moran J, Pawlicki T, Molineu A, et al. Tolerance limits and methodologies for IMRT measurement-based verification QA: recommendations of AAPM Task Group N0. 218. Med Phys 2018;45(4):e53–83.
[32] Galloway RMM. Texture analysis using gray level run lengths. Comput Graph Image Process 1975;4:172–9.
[33] Qadri S, Khan DM, Ahmad F, Qadri SF, Babar ME, Shahid M, et al. A comparative study of land cover classification by using multispectral and texture data. Biomed Res Int 2016;8797438:1–12.
[34] Conners RW, Harlow CA. A theoretical comparison of texture algorithms. IEEE Trans Pattern Anal Mach Intell 1980;2:204–22.
[35] Leijenaar RT, Carvalho S, Velazquez ER, van Elmpt WJ, Parmar C, Hoekstra OS, et al. Stability of FDG-PET Radiomics features: an integrated analysis of test-retest and inter-observer variability. Acta Oncol 2013;52:1391–7.
[36] Garcia G, Maiora J, Tapia A, De Blas M. Evaluation of texture for classification of abdominal aortic aneurysm after endovascular repair. J Digit Imaging 2012;25:369–76.
[37] Mir AH, Hanmandlu M, Tandon SN. Texture analysis of CT-images for early detection of liver malignancy. Biomed Sci Instrum 1995;31:213–7.
[38] Kim JK, Park HW. Statistical textural features for detection of microcalcifications in digitized mammograms. IEEE Trans Med Imaging 1999;18:231–8.