scientific reports

OPEN

Check for updates

Application of error classification model using indices based on dose distribution for characteristics evaluation of multileaf collimator position errors

Heesoon Sheen^{1,5}, Han-Back Shin^{2,3,5}, Hojae Kim⁴, Changhwan Kim³, Jihun Kim³, Jin Sung Kim³ & Chae-Seon Hong^{3⊠}

This study aims to evaluate the specific characteristics of various multileaf collimator (MLC) position errors that are correlated with the indices using dose distribution. The dose distribution was investigated using the gamma, structural similarity, and dosiomics indices. Cases from the American Association of Physicists in Medicine Task Group 119 were planned, and systematic and random MLC position errors were simulated. The indices were obtained from distribution maps and statistically significant indices were selected. The final model was determined when all values of the area under the curve, accuracy, precision, sensitivity, and specificity were higher than 0.8 (p < 0.05). The dose–volume histogram (DVH) relative percentage difference between the error-free and error datasets was examined to investigate clinical relations. Seven multivariate predictive models were finalized. The common significant dosiomics indices (GLCM Energy and GLRLM_LRHGE) can characterize the MLC position error. In addition, the finalized logistic regression model for MLC position error prediction showed excellent performance with AUC > 0.9. Furthermore, the results of the DVH were related to dosiomics analysis in that it reflects the characteristics of the MLC position error. It was also shown that dosiomics analysis could provide important information on localized dose-distribution differences in addition to DVH information.

Intensity-modulated radiation therapy (IMRT) enables highly conformal and precise dose distribution to the target, reducing the exposure of healthy tissue from unwanted radiation^{1,2}. Such remarkable therapeutic outcomes are achieved using a unique beam-shaping device called the multileaf collimator (MLC)^{1,3–6}. It is a crucial component of the modulation plan defined for IMRT to generate complex dose distributions⁷. Therefore, regularly monitoring MLC's accuracy and reproducibility are important to ensure that the actual and planned locations match during treatment^{8–10}.

The issue of MLC position uncertainty has been widely investigated for different systems and techniques using radiographic film, ionization chambers, portal imaging devices, and log files¹¹⁻¹⁸. However, the utilization of different detectors may yield different results because of their unique characteristics (resolution, detector density, geometric (2D/3D) differences, and calibration method) and potential differences in the gamma calculation method¹⁸⁻²⁰.

Gamma analysis^{21–23}, developed to quantitatively evaluate the consistency of dose distributions, has been implemented using a combined dose difference (DD) and distance to agreement (DTA) threshold to determine whether each IMRT plan is acceptable in regular clinical practice. However, several groups^{4,7,23} have reported the following limitations in evaluating the uncertainty related to the MLC leaf position: (1) insensitivity to dose errors, (2) absence of correlations with clinical dose errors, (3) difficulty in defining the root cause of the existing discrepancy between dose distributions, (4) differences between the gamma results produced by different

¹Department of Health Sciences and Technology, Samsung Advanced Institute for Health Sciences and Technology, Sungkyunkwan University, Seoul, South Korea. ²Department of Radiation Oncology, Gachon University Gil Medical Center, Incheon, South Korea. ³Department of Radiation Oncology, Yonsei Cancer Center, Yonsei University College of Medicine, Seoul, South Korea. ⁴Department of Radiation Oncology, Yonsei Cancer Center, Seoul, South Korea. ⁵These authors contributed equally: Heesoon Sheen and Han-Back Shin. ^[2]email: cs.hong@yuhs.ac measurement devices²⁴, and (v) different standard criteria proposed in the Imaging and Radiation Oncology Core (IROC) and TG218 reports²⁵. Recently, dosiomics (or radiomics)^{1,4,7,26–29} and structural similarity (SSIM)⁴ have been conducted to determine whether they can be used to overcome the limitations of conventional gamma analysis. First, dosiomics (or radiomics) analysis has been applied to extract valuable features from dose maps between the planned and delivered doses in patient-specific quality assurance (PSQA) by considering the dose or fluence map as an image. Several studies have shown that radiomics-based analysis can effectively detect MLC positioning errors and provide an alternative to conventional gamma analysis^{4,5,11,26}. Second, the SSIM index can detect absolute dose errors, gradient discrepancies, and dose structure errors with sub-indices luminance, contrast, and structure^{1,4,7,25–27}. Both methods can be employed to describe MLC errors because they show the locations of large discrepancies and different types of error-related patterns. However, no strong indicators or harmonized error characterization or prediction results have been reported because of measurement device bias and patient-specific heterogeneity when using patient data^{1,4,7,25–29}. For these reasons, essential research is needed to discover and analyze basic indicators in a state where heterogeneity due to measurement devices and patients is removed before direct application to the clinical practice of various conditions^{1,4,7,26–34}.

This study aims to evaluate the characteristics of various MLC position errors correlated with indices using dose distribution, and to use them as reference data for future clinical applications. In this regard, two implementations were preferentially performed as follows: to exclude clinical variability, a single-material AAPM TG-119 phantom was used to establish a treatment plan; MLC position errors were generated in the treatment planning system (TPS) instead of measurements. The difference between the error-free and error dose distributions was analyzed using the gamma index, SSIM index, and dosimetry index to investigate the characteristics of the MLC position error prediction model. The dose-volume histogram (DVH) was examined to investigate the clinical relationship. We evaluated the clinical significance by analyzing the final optimal model and DVH parameter results. To the best of our knowledge, this study is the first attempt to analyze the characteristics and predictive models of systematic and random MLC error using dosiomics, gamma, and SSIM indices in the absence of deviation due to clinical, instrumental, and technical differences.

Materials and methods

IMRT treatment plans for American Association of Physicists in Medicine³⁵ **Task Group 119 (TG-119) cases.** All plan data were generated and optimized using the segmental MLC (SMLC) mode in the RayStation planning system v5 (RaySearch Laboratories, Stockholm, Sweden). Computed tomography (CT) and structure datasets including phantoms used in this study were downloaded from AAPM TG-119 case³⁶. Additionally, AAPM TG-119 suggests the IMRT goals and beam arrangement. For each plan, beam arrangements and planned doses, such as angle and number of beam field on Elekta Versa HD linear accelerator with Agility MLC (Elekta AB, Stockholm, Sweden), were set as recommended by AAPM TG 119 (Supplementary Table S1). The plans were generated using RayStation v5 TPS with the dose calculation algorithm (collapsed cone convolution (CCC)) of a grid size of 2.0 mm for each beam of the static IMRT plans using segmented MLC.

Simulated MLC leaf position error. This study primarily considered two types of MLC positional errors: systematic error and random error. DICOM-RT plan files were extracted from RayStation v5 to simulate the MLC positional errors. The original MLC position for all the control points described in the DICOM-RT plan file was changed to the locations specified in the DICOM-RT plan file and modified using in-house software(see Supplementary MATLAB source code), developed using MATLAB 2018b (Mathworks, Natick, MA). Subsequently, the modified DICOM-RT plan files were imported back into RayStation v5, as shown in Fig. 1 ³⁶.To evaluate only the effect of the MLC position error, the MU values on all control points between the error-free plan and error plan were not changed, and the same plan objective function was used. Moreover, only recalculation was performed without optimization based on the CCC algorithm.

As shown in Fig. 1, the two types of MLC position errors were induced in each beam of the error-free plan as follows:

For the systematic MLC position error plans, at every control point, the MLC positions surrounding the planning target volume (PTV) were shifted by 0.5 mm, 1.0 mm, 1.5 mm, and 2.0 mm to the right of their original leaf position on one side of the bank.

For the random MLC position error plans, at every control point, the MLC positions in both banks were randomly shifted by a pseudo-random number having a Gaussian distribution of mean value, μ , 0.0 mm, 1.0 mm, and 2.0 mm, with a standard deviation, σ , 1.0 mm width (1 sigma)^{10,37}. When the shift of the leaf position collided with the leaf located on the opposite side bank, it was randomly arranged within the corresponding Gaussian distribution to avoid a collision.

We developed 35 simulated MLC position error plans and total cases for each angle and beam (Table 1).

Gamma analysis. The gamma-index method evaluates the coincidence between two dose distributions of error-free and error datasets using the percent DD and DTA³⁸. We performed a complete local 3D gamma analysis between two dose files for the error-free and error plans with error based on the criteria 3%/3 mm, 3%/2 mm (the AAPM TG-218 recommendation), 2%/2 mm, and 1%/1 mm with a 10% threshold using PTW Verisoft software, version 6.1 (PTW, Freiburg, Germany)^{7,10,38,39}. Three-dimensional gamma analysis was conducted for each angle and the entire beam.

SSIM. The SSIM index is designed to compare and evaluate pairs of images (error-free and error-induced plan) and can be utilized to evaluate the luminance, contrast, and structure of SSIM^{4,25,38,40}.



Figure 1. Schematic workflow of simulated systematic and random MLC errors.

Number of beams	Head and neck	Prostate	C-shape
Error-free	36	28	36
Systematic error	36	28	36
Random error	27	21	27
Total	99	77	99

Table 1. MLC position-error types and numbers.

In this study, the entire SSIM index and three subcomponents (luminance, contrast, and structural index) were also evaluated as functions of the beam fields. Preprocessing was not required for SSIM because the size of dose maps was the same. The SSIM index was calculated using MATLAB 2018b (Mathworks, Natick, MA), and it is expressed as follows:

$$SSIM(x,y) = [l(x,y)]^{\alpha} \cdot [c(x,y)]^{\beta} \cdot [s(x,y)]^{\gamma}.$$
(1)

The default SSIM index is based on the following settings: $\alpha = \beta = \gamma = 1$, and $C_3 = \frac{C_2}{2}$, and the SSIM index can be calculated as follows:

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)},$$
(2)

where l(x, y), c(x, y), and s(x, y) are the luminance, contrast, and structure subindex, respectively. μ_x and μ_y , σ_x and σ_y , and σ_{xy} are the local means, standard deviations, and cross-variance for images x and y, respectively.

The specific parameters for the SSIM calculation were set based on previous studies^{4,25,40}. The regularization constant is calculated as $C_1 = (K_1L)^2$, $C_2 = (K_2L)^2$, and $C_3 = C_2/2$. K_1 and K_2 were set to 0.01 and 0.03, respectively, as the default values suggested by Wang et al.⁴⁰. Peng et al.²⁵ proposed the default values of K_1 and K_2 as suitable factors for evaluating the MLC position error in the result of the regularization constant effect according to the K_1 and K_2 values. The dynamic range, L, was set to 200, corresponding to the fraction dose in this study.

Dosiomics analysis. For dosiomics analysis, two different dose-distribution datasets were generated: (1) subtracted error-free datasets (simulated errorfree dose map—errorfree dose map) and (2) subtracted error datasets (simulated error dose map—error-free dose map) (Fig. 2). The error-free dose map represents the dose distribution extracted after planning for all control points at the original MLC location via RayStation. To generate a subtracted error-free dose map for dosiomics analysis, two error-free dose maps were required. If the same-dataset error-free dose map (unmodified dose map) is used as a simulated error-free dose map to create sub-error-free dose map was generated by systematically moving the error-free dose map by 0.01 mm. The similarity and correlation between simulated error-free and unmodified error-free dose maps were examined using Wil-



Figure 2. Dose maps. (a) Error-free dose map (b) Error-induced dose map (c) Subtracted (error-induced – error-free) dose map.

coxon signed-rank test and Spearman's rank correlation. Their results showed statistically significant similarity (*p* value > 0.05) in Wilcoxon signed-rank test and the strong correlation (coefficient > 0.97 and *p* value < 0.001) in Spearman's rank correlation (Supplementary Table S2). Therefore, it was confirmed that the simulated error-free had similarity enough to generate sub-error-free when used to run in this experiment. The subtracted dose maps were classified into three types: sub-error-free (simulated error-free dose map—error-free dose map), sub-systematic-error (simulated systematic error dose map—error-free dose map), and sub-random-error (simulated random error dose map—error-free dose map). Class-I included the error-free type and combined systematic-error free and random-error types. Error-type classes in subtracted dose maps are summarized in Table 2.

This study performed the dosiomics analysis for Class-I, Class-II, and Class-III. The 275 fluence maps of the four treatment plans exported in the DICOM-RT file from the TPS were analyzed using the Local Image Features Extraction (LIFEx) version 7.1.0 software package (http://www.lifexsoft.org)⁴¹. For dosiomics index calculations, spatial resampling was 2 mm (X-direction), 2 mm (Y-direction), and 2 mm (Z-direction) in Cartesian coordinates. The size of the bin in intensity discretization was 1. Thirty-four radiomics features were categorized into one conventional feature, two histogram features, and thirty-one texture features^{42–44}. Four matrices, namely CLCM, CLRLM, NGLDM, and GLZLM^{26,42–44} were used to determine thirty-one texture features. GLCM was obtained in 13 directions in 3D with one voxel distance relationship between neighboring voxels to indicate the arrangements of pairs of voxels used to calculate textural features. GLRLM was calculated for the 13 different directions to represent the size of homogeneous runs for each gray level. NGLDM was related to the difference in gray levels between one voxel and its 26 neighbors in 3D. GLZLM was calculated directly in 3D to explain the size of homogeneous zones for each gray level. The radiomics features were extracted from the whole subtracted dose map and standardized them to obtain the standard score (z-score) (Supplementary Table S3).

Statistical analysis. All statistical analyses were performed using RStudio (version 2021.09.1-372; RStudio Software Inc. (Boston, MA, USA)). The error-free data were labeled as "0," and systematic error data and random error data were labeled as "1." Gamma, SSIM, and dosiomics indices were examined and selected for developing the MLC position error prediction model. The independence of all the indices of the gamma, SSIM, and dosiomics was investigated to prevent overfitting using Spearman's rank correlation, backward stepwise elimination, and multicollinearity. First, the indices with Spearman's rank correlation coefficient higher than 0.8 were removed after the Holm–Bonferroni correction method was applied for all *p* values to correct multiple test comparisons. Then, the remaining indices were filtered by performing backward stepwise elimination. The indices selected through these two steps were selected by the multicollinearity using the variance inflation factor (VIF < 4)⁴⁵. In addition, univariate and multivariate logistic regression were also used for index selection.

Predictive model development and performance. Univariate and multivariate logistic regression models for MLC position error prediction was built in RStudio through the following process. To ensure repro-

Class of error	Errors	
Class-I	Sub-error-free	
	Sub-systematic-error + Sub-random-error	
Class-II	Sub-error-free	
	Sub-systematic-error	
Class-III	Sub-error-free	
	Sub-random-error	

Table 2. Error-type classes in subtracted dose map.

ducibility of the random sampling, the '*set.seed*' function was implemented. The dataset was loaded and split into training and testing sets using the '*createDataPartition*' function. The numbers of training datasets corresponded to 60%–80% of the total number datasets (sub-error-free datasets: 36, sub-systematic-error datasets: 36, sub-random-error datasets: 27) (Supplementary Table S4). The '*trainControl*' function was used to define the training control object, which specified that tenfold cross-validation should be used, repeated 10 times using '*repeatedcv*' method. The '*twoClassSummary*' function was used to summarize the results of the cross-validation process, and '*classProbs*' was set to TRUE to enable probability estimates. The '*grid*' was used to search for the best hyperparameters in the tuning grid. The '*train*' function was then used to train a logistic regression model on the training data using the '*glm*' method and the training control object. The '*Accuracy*' metric was used to evaluate the accuracy of a classification model in the train function. The '*predict*' function was used to predict the class labels for the test data, and the '*confusionMatrix*' and '*roc*' functions were used to evaluate the model in terms of sensitivity, specificity, accuracy, and precision, the area under the curve (AUC) computed based on receiver operating characteristic curve (ROC). The workflow for developing the prediction error model is illustrated in Fig. 3.

DVH. DVHs for AAPM TG-119³⁶ cases with planned IMRT treatment were generated using RayStation v5. The criteria for meaningful differences were underdosed for the PTV and 3% or more for the OAR based on studies on site-based patient quality assurance (QA) standards^{46–48}.

Results

Gamma and SSIM. The gamma index was analyzed between the error-free and induced error plans, and each documented angle and the total dose in a 3D dose file were analyzed. For the head and neck cases, the gamma-index ranges were 0.926–1.0, 0.848–1.0, 0.758–1.0, and 0.463–0.946 for the 3 mm/%, 2 mm/3%, 2 mm/2%, and 1 mm/1% criteria, respectively. The average SSIM, luminance, contrast, and structure indices were 0.9023–0.9534, 0.9926–0.9995, 0.9983–1.000, and 0.9942–0.9999, respectively. The head and neck cases are summarized in Table 3. The other cases are listed in Supplementary Table S5 (Fig. 4).

Dosiomics analysis. The dosiomics indices were extracted from the 275 subtracted dose maps (Class-I, Class-II, and Class-III) generated by all beam fields of the four IMRT plans. Specifically, 34 dosiomics indices were selected through Spearman's rank correlation, backward stepwise elimination, and VIF in all cases. Figure 5 illustrates the number of features according to the plan type. The gray-level run length matrix (GLRLMs) accounted for 35.8% of all features (19/53), and GLCMs approached approximately 28.3%. Eleven, 11, 16, and 15 dosiomics indices were selected from the head and neck, prostate, C-shape easy, and C-shape hard, respectively. The gray-level co-occurrence matrix energy (GLCM_Energy) was the most common dosiomics indices in Class-I and Class-III. Gray-level run length matrix long run high gray-level emphasis (GLRLM_LRHGE) was Class-II's most common dosiomics index.



Figure 3. Workflow for developing predictive error models.

Head and neck	Gamma index	SSIM index	Luminance	Contrast	Structure		
Systematic error							
0.5	0.9999	0.9366	0.9982	0.9997	0.9991		
1.0	0.9948	0.9360	0.9982	0.9997	0.9991		
1.5	0.9591	0.9346	0.9981	0.9997	0.9991		
2.0	0.9049	0.9323	0.9980	0.9997	0.9990		
Random error							
0.0	1.0000	0.9361	0.9982	0.9997	0.9991		
0.5	0.9972	0.9362	0.9982	0.9997	0.9991		
1.0	0.9883	0.9357	0.9982	0.9997	0.9991		

Table 3. Mean gamma index (2 mm/3%), mean SSIM index, and subcomponent (luminance, contrast, structure) indices for the head and neck cases.







Figure 5. Stacked bar plot showing the distribution of selected features as a function of plan type, with data labels showing the number of features.

Predictive model development and performance. The independent indices of gamma, SSIM, and dosiomics were used to develop predictive models for detecting the MLC position errors after indices were excluded using Spearman's rank correlation (coefficient > 0.8 and p-value < 0.05), backward stepwise elimination, and VIF (≥ 4) to avoid overfitting among the indices. The independent indices selected using Spearman's rank correlation and confusion matrix are shown in Supplementary Table S6.

In the final prediction model, only dosiomics indices were selected as significant indices that satisfy the statistical criteria. Therefore, only dosiomics indices were used to develop a dose distribution-based MLC error

prediction model. The developed models were evaluated using an independent test dataset. Consequently, the final models were determined when the model's accuracy, precision, sensitivity, and specificity were 0.8 or more (p < 0.05) in Class-I, Class-II, and Class-III. In Class-I, five univariate predictive models (one model each for the head and neck, one for prostate cases, two for the C-shape easy cases, and one for C-shape hard cases) were decided as the final models (Supplementary Table S7). In Class-II, five univariate final predictive models (two models each for the head and neck, one for the prostate, and C-shape easy and hard cases) were determined (Supplementary Table S7). In Class-III, four univariate models were finalized (one for head and neck, prostate, c-shape easy, and C-shape hard cases) (Supplementary Table S7). Six multivariate models were also developed (Supplementary Table S7).

DVH analysis. The DVHs of the error-free datasets were compared to those of the four systematic error datasets and three random error datasets in terms of clinical effectiveness. In the head and neck case, the relative percentage differences of Sys 1.5 mm, Sys 2.0 mm, and Ran 1.0 mm were greater than 3.0% compared with the DVH of the error-free cases. This indicates that the dose received by the cord and parotids was 3.0% higher. In the prostate case, the dose received by the bladder above the PTV was greater than 3.0% for Sys 1.5 mm, Sys 2.0 mm, and Ran 1.0 mm. The rectum below the PTV received a 3% higher dose than in the error-free case for Sys 1.5 mm, Sys 2.0 mm, and Ran 1.0 mm. In the two C-shape cases, the dose of the cord was 3.0% higher than those in the error-free cases for all error types. The DVH relative percentage difference was almost linear in the systematic error. However, the DVH relative percentage difference of random error showed different tendencies depending on the location of the structure (Supplementary Fig. S1).

Discussion

In this study, the characteristics of the MLC position error under non-heterogeneous conditions were explained using only the dosiomics indices because they were statistically significant among the gamma, SSIM, and dosiomics indices. The determined dosiomics indices were used for the predictive MLC position error model. The clinical relationship of significant indices and prediction MLC position error model was examined using DVH.

For gamma-index results, even in the case of MLC position systematic 1.0 mm shift, it calculated 0.905 for the C-shape easy case and 0.710 for the C-shape hard case. These results indicate that the dose distribution was affected by plan complexity, and the gamma index was considered to be low. However, there was no significant difference within 3% in the DVH index for PTV and cord. The gamma index had a low correlation with the DVH parameter. The SSIM-index tendency according to the random MLC displacement was relatively small and irregular compared with the SSIM index according to the systematic MLC displacement (Table 3, Fig. 4, and Supplementary Table S5). For the systematic error, only the luminance SSIM subcomponent was more sensitive than the other SSIM subcomponents for the MLC position error. However, its sensitivity was relatively small compared with the dosiomics indices (Supplementary Table S5). For the random error, the gamma and SSIM indices did not convey any trend of the MLC position error.

Among the gamma, SSIM, and dosiomics indices, statistically significant indices representing the characteristics of the MLC position error were extracted from dosiomics. In Class-I, GLCM_Energy was selected as the common significant index for all predictive models. It belongs to the GLCM, representing the dose distribution with co-occurring pixel values at one offset. It was showed that increased error increased the difference between error-free and erroneous data and thus increased non-uniformity of gray level voxel pairs. In Class-II, GLRLM_LRHGE was selected as the common significant index for all cases, and it belongs to the GLRLM, representing the size of the homogenous run. GLRLM_LRHGE shows the distributions with long homogenous runs with high gray levels. As a result, the feature of long homogenous runs with high gray levels was presented for the MLC systematic position error. In Class-III, GLCM_Energy was chosen as the common significant dosiomics index, as in Class-I, except for C-shape hard cases where the significant index was gray-level zone length matrix (GLZLM) gray level_nonuniformity (GLNU). Our result that more than half of the statistically significant indices belonged to the GLCMs was consistent with that of Ma et al.⁴. In addition, GLRLM_LRHGE that can detect systematic errors, and GLCM_Energy that can detect nonuniformity or random errors, are consistent with the results of the paper published by Landon S. Wootton et al.²⁶. These two studies were performed using different devices and techniques than ours, and while our study did not include clinical variations, these two studies did include them. Nevertheless, the significant dosiomics indices we found were consistent with those in these studies. These consistent results confirmed that the significant indices we found were a basic index that characterizes the MLC position error regardless of the measurement device, technique, and clinical variation.

The predictive models for the MLC position errors were developed using only dosiomics indices in Class-I, Class-II, and Class-III. The gamma and SSIM indices were disregarded because they were not dedicated adequate weights for developing the predictive models, compared to the dosiomics index. The final error prediction models using the significant dosiomics indices describing the characteristics of the MLC position error exhibited excellent performance with AUC>0.9, and accuracy, sensitivity, and specificity \geq 0.8 (p<0.05), except for the C-shape hard case of Class-III. The more the plan complexity, the more distributed was the influence of the MLC position error; therefore, it was estimated that the accuracy, precision, and specificity values other than AUC and sensitivity values were less than 0.8.

In the DVH analysis, the relative percentage difference in systematic error was observed to increase almost linearly. This indicates that the larger the systematic error in which the MLC position is offset in one direction, the larger is the area where the dose distribution differs. For a random error in which the MLC position is offset in a random direction, the DVH analysis result showed that the relative percentage error increased with different trends depending on the structure's location. The reason was that the structures in different locations were affected differently for random errors because MLC position errors occur in a random direction. As a result, the

systematic error affected the size of the homogenous run in the dose-distribution difference. By contrast, the random error affected the voxel pairs in the dose-distribution difference. These results confirmed that the common significance index indicating the characteristics of the differential dose distribution for systematic errors was GLRLM_LRHGE, and that of random errors was related to GLCM_Energy. Regarding the complexity of the plan, the DVH result showed that the more complex the plan, the greater the relative percentage difference. However, no common significant index representing the dose-distribution difference was found. This is probably because the greater scatter of nonuniformity showing a significant common index as the complexity of the plan increases. These characteristics make different dose distributions depending on the plan complexity, and the significant dosiomics index also varies owing to the resulting random dose-distribution differences. This phenomenon was confirmed in the examples of the C-shape easy case and C-shape hard case. Each of them had two different significant indices, GLCM_Energy and GLZM_GLNU. While GLCM_Energy selected from the C-shape easy case considers a pixel pair with a specific value, the significant index GLZM_GLNU chosen from the C-shape hard case considers the connected voxels. Therefore, depending on the complexity of the plan, the DVH provides information on the dose-volume histogram but cannot show the localized texture differences between dose distributions. In the case of dosiomics analysis, it can provide essential additional information that DVH cannot offer because it can show differences in the local texture of the dose distribution depending on the complexity of the plan.

The results of our study are summarized in three key findings as follows. First, the discovery of principle common significance indices (GLCM_Energy, GLRLM_LRHGE) represents the characteristics of systematic and random MLC position errors and can be used as reference indices for future clinical applications. Second, the prediction MLC position error models developed using statistically significant dosiomics indices showed superior performance (AUC>0.9). Predictive models have been developed based on specific clinical sites. However, given the significant common indicators found in this study, it is necessary to investigate whether models generated in mixed clinical settings can exhibit excellent predictive power through future studies. Third, the results of DVH in the primary state were confirmed to be related to dosiomics analysis, which represented the characteristics of the dose-distribution difference due to the MLC position error. In addition, it has been demonstrated that in cases with different plan complexity, dosiomics analysis can give more critical information that can be added to the information in the DVH. The results of our study can ultimately provide information related to the characteristics of localized texture using dosiomics analysis. This information could be made more clinically useful by providing additional information to the gamma index and DVH commonly provided by devices and technologies such as log file-based QA or electronic portable imaging devices for dosimetric impact assessment.

Although our results demonstrated the novel clinical findings mentioned above, our study has four limitations. First, the dosiomics index was extracted from the subtracted dose distribution but the gamma and SSIM indices were extracted from dose distributions of the error-free and simulated error instead of their subtracted dose distribution. Second, because only one type of error (MLC position error) was studied under the condition that the factors affecting the error analysis are removed, the direct application of the result to various clinical processes is limited. Third, understanding the relationship between the grid size, preprocessing filter and MLC position error is necessary. In this study, the characterization analysis for the MLC position error uncertainty smaller than the grid size was established by referring to other studies^{7,49,50}. Although the MLC position error was smaller than the grid size, there was a difference in the dose distribution; however, the effect of the MLC position error was not fully reflected. Therefore, it is necessary to conduct a study considering the relationship between the grid size and MLC position error. Furthermore, filtered features such as wavelets and Laplacian-of-Gaussian filters were not used as they could potentially introduce noise and artifacts into the dose map which can affect the accuracy and reliability of the results⁵¹. However, these filters play important roles in dosiomics by emphasizing specific image characteristics, such as edges, and extracting biomarkers^{2,51,52}. Therefore, further research is needed to investigate the use of these filters in dosiomics and to develop appropriate methods to mitigate the potential introduction of noise and artifacts. In addition, the features used in this study provide information about the spatial and statistical distribution of dose levels and have been considered sufficient for estimating the MLC error from subtractive dose maps. On the other hand, other studies^{1,4} showed the predictive ability using more dosiomic features. Therefore, it needs to investigate the correlation between the number of dosiomic features and the predictive performance of radiomic models in future studies. Forth, this study used the plan results of static IMRT using segmented MLC to evaluate the characteristics of MLC. However, additional research is needed on how it is reflected when volumetric modulated arc therapy (VMAT) and dynamic MLC, which are clinically applied plan techniques, are applied. In the future, we plan to investigate the detection sensitivity utilizing dosiomics, SSIM, and DVH according to diverse error scenarios and complexities in heterogeneous environments for clinical application.

Conclusion

The characteristics of the MLC position error were investigated using dose-distribution comparisons of errorfree and simulated error datasets. Predictive models were developed and evaluated for the accurate and precise analysis of the MLC position error. Our study highlights three novel results. First, MLC position-error characteristics were described using common significant dosiomics indices (GLCM_Energy, and GLRLM_LRHGE). Second, the finalized logistic regression model for MLC position error prediction showed excellent performance with AUC>0.9. Third, it was confirmed that the effects of MLC position error on DVH are related to the dosiomics analysis and that the dosiomics analysis provides the important information on localized texture of dose distribution in addition to the DVH information. These clinically significant results are expected to be used as primary data to discover and study clinically meaningful indicators.

Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Received: 13 September 2022; Accepted: 20 May 2023 Published online: 07 July 2023

References

- Lizar, J. C., Yaly, C. C., Colello Bruno, A., Viani, G. A. & Pavoni, J. F. Patient-specific IMRT QA verification using machine learning and gamma radiomics. *Phys. Med.* 82, 100–108. https://doi.org/10.1016/j.ejmp.2021.01.071 (2021).
- Van Griethuysen, J. J. M. et al. Computational radiomics system to decode the radiographic phenotype. Cancer Res. 77, e104–e107. https://doi.org/10.1158/0008-5472.CAN-17-0339 (2017).
- Taylor, A. & Powell, M. E. Intensity-modulated radiotherapy—what is it?. Cancer Imaging 4, 68–73. https://doi.org/10.1102/1470-7330.2004.0003 (2004).
- 4. Ma, C. *et al.* The structural similarity index for IMRT quality assurance: Radiomics-based error classification. *Med. Phys.* **48**, 80–93. https://doi.org/10.1002/mp.14559 (2021).
- 5. Potter, N. J. *et al.* Error detection and classification in patient-specific IMRT QA with dual neural networks. *Med. Phys.* 47, 4711–4720. https://doi.org/10.1002/mp.14416 (2020).
- Hong, C. S. *et al.* Carotid-sparing TomoHelical 3-dimensional conformal radiotherapy for early glottic cancer. *Cancer Res. Treat.* 48, 63–70. https://doi.org/10.4143/crt.2014.265 (2016).
- Sakai, M. et al. Detecting MLC modeling errors using radiomics-based machine learning in patient-specific QA with an EPID for intensity-modulated radiation therapy. Med. Phys. 48, 991–1002. https://doi.org/10.1002/mp.14699 (2021).
- Agnew, A., Agnew, C. E., Grattan, M. W., Hounsell, A. R. & McGarry, C. K. Monitoring daily MLC positional errors using trajectory log files and EPID measurements for IMRT and VMAT deliveries. *Phys. Med. Biol.* 59, N49-63. https://doi.org/10.1088/0031-9155/ 59/9/n49 (2014).
- Woo, M. K. & Nico, A. Impact of multileaf collimator leaf positioning accuracy on intensity modulation radiation therapy quality assurance ion chamber measurements. *Med. Phys.* 32, 1440–1445. https://doi.org/10.1118/1.1901843 (2005).
- Yan, G. et al. On the sensitivity of patient-specific IMRT QA to MLC positioning errors. J. Appl. Clin. Med. Phys. 10, 120–128. https://doi.org/10.1120/jacmp.v10i1.2915 (2009).
- Chuang, K. C., Giles, W. & Adamson, J. A tool for patient-specific prediction of delivery discrepancies in machine parameters using trajectory log files. *Med. Phys.* 48, 978–990. https://doi.org/10.1002/mp.14670 (2021).
- Stevens, M. T. R., Lobb, E. C. & Yenice, K. M. Validation of MLC-based linac radiosurgery for trigeminal neuralgia. J. Appl. Clin. Med. Phys. 19, 214–221. https://doi.org/10.1002/acm2.12381 (2018).
- Li, Y., Chen, L., Zhu, J., Wang, B. & Liu, X. A quantitative method to the analysis of MLC leaf position and speed based on EPID and EBT3 film for dynamic IMRT treatment with different types of MLC. J. Appl. Clin. Med. Phys. 18, 106–115. https://doi.org/ 10.1002/acm2.12102 (2017).
- Visser, R. et al. Reconstruction of high resolution MLC leaf positions using a low resolution detector for accurate 3D dose reconstruction in IMRT. Phys. Med. Biol. 61, N642-n649. https://doi.org/10.1088/0031-9155/61/23/n642 (2016).
- Chen, A., Zhu, J., Wang, N., Chen, L. & Chen, L. Comparison of three-dimensional patient-specific dosimetry systems with delivery errors: Toward a new synchronous measurement method. *Physica Med.* **90**, 134–141. https://doi.org/10.1016/j.ejmp.2021.09.013 (2021).
- Olasolo-Alonso, J., Vázquez-Galiñanes, A., Pellejero-Pellejero, S. & Pérez-Azorín, J. F. Evaluation of MLC performance in VMAT and dynamic IMRT by log file analysis. *Physica Med.* 33, 87–94. https://doi.org/10.1016/j.ejmp.2016.12.013 (2017).
- Kabat, C. N. et al. Evaluation of the Elekta agility MLC performance using high-resolution log files. Med. Phys. 46, 1397–1407. https://doi.org/10.1002/mp.13374 (2019).
- Gray, A. et al. Evaluation of the ability of three commercially available dosimeters to detect systematic delivery errors in step-andshoot IMRT plans. Rep. Pract. Oncol. Radiother. 26, 793–803. https://doi.org/10.5603/RPOR.a2021.0093 (2021).
- Alharthi, T., Vial, P., Holloway, L. & Thwaites, D. Intrinsic detector sensitivity analysis as a tool to characterize ArcCHECK and EPID sensitivity to variations in delivery for lung SBRT VMAT plans. J. Appl. Clin. Med. Phys. 22, 229–240. https://doi.org/10. 1002/acm2.13221 (2021).
- Hong, C. S. et al. Comparison of the Efficacy of 2D Dosimetry Systems in the Pre-treatment Verification of IMRT. J. Korean Soc. Ther. Radiol. Oncol. 27(2), 91–102. https://doi.org/10.3857/jkstro.2009.27.2.91 (2009).
- 21. Low, D. A., Moran, J. M., Dempsey, J. F., Dong, L. & Oldham, M. Dosimetry tools and techniques for IMRT. *Med. Phys.* 38, 1313–1338. https://doi.org/10.1118/1.3514120 (2011).
- Depuydt, T., Van Esch, A. & Huyskens, D. P. A quantitative evaluation of IMRT dose distributions: Refinement and clinical assessment of the gamma evaluation. *Radiother. Oncol.* 62, 309–319. https://doi.org/10.1016/s0167-8140(01)00497-2 (2002).
- Baran, M. et al. Are gamma passing rate and dose-volume histogram QA metrics correlated?. Med. Phys. 48, 4743–4753. https:// doi.org/10.1002/mp.15142 (2021).
- Feygelman, V., Zhang, G., Stevens, C. & Nelms, B. E. Evaluation of a new VMAT QA device, or the "X" and "O" array geometries. J. Appl. Clin. Med. Phys. 12, 3346. https://doi.org/10.1120/jacmp.v12i2.3346 (2011).
- Peng, J. et al. Implementation of the structural SIMilarity (SSIM) index as a quantitative evaluation tool for dose distribution error detection. Med. Phys. 47, 1907–1919. https://doi.org/10.1002/mp.14010 (2020).
- Wootton, L. S., Nyflot, M. J., Chaovalitwongse, W. A. & Ford, E. Error detection in intensity-modulated radiation therapy quality assurance using radiomic analysis of gamma distributions. *Int. J. Radiat. Oncol. Biol. Phys.* 102, 219–228. https://doi.org/10.1016/j. ijrobp.2018.05.033 (2018).
- Nyflot, M. J., Thammasorn, P., Wootton, L. S., Ford, E. C. & Chaovalitwongse, W. A. Deep learning for patient-specific quality assurance: Identifying errors in radiotherapy delivery by radiomic analysis of gamma images with convolutional neural networks. *Med. Phys.* 46, 456–464. https://doi.org/10.1002/mp.13338 (2019).
- Park, S. Y., Kim, J. I., Hoon Oh, D. & Park, J. M. Evaluation of the plan delivery accuracy of intensity-modulated radiation therapy by texture analysis using fluence maps. *Physica Med.* 59, 64–74. https://doi.org/10.1016/j.ejmp.2019.02.016 (2019).
- Placidi, L. et al. Stability of dosomics features extraction on grid resolution and algorithm for radiotherapy dose calculation. Physica Med. 77, 30–35. https://doi.org/10.1016/j.ejmp.2020.07.022 (2020).
- Liang, B. et al. Dosiomics: Extracting 3D spatial features from dose distribution to predict incidence of radiation pneumonitis. Front. Oncol. 9, 269. https://doi.org/10.3389/fonc.2019.00269 (2019).
- 31. Buizza, G. *et al.* Radiomics and dosiomics for predicting local control after carbon-ion radiotherapy in skull-base chordoma. *Cancers* 13, 1–15. https://doi.org/10.3390/cancers13020339 (2021).
- Placidi, L. et al. A multicentre evaluation of dosiomics features reproducibility, stability and sensitivity. Cancers https://doi.org/ 10.3390/cancers13153835 (2021).

- Hirashima, H. *et al.* Improvement of prediction and classification performance for gamma passing rate by using plan complexity and dosiomics features. *Radiother. Oncol.* 153, 250–257. https://doi.org/10.1016/j.radonc.2020.07.031 (2020).
- Kadoya, N. *et al.* Quantifying the performance of two different types of commercial software programs for 3D patient dose reconstruction for prostate cancer patients: Machine log files vs. machine log files with EPID images. *Physica Med.* 45, 170–176. https:// doi.org/10.1016/j.ejmp.2017.12.018 (2018).
- AAPM. TG-119 IMRT commissioning tests instructions for planning, measurement, and analysis. *Differences* 21, 308–310 (2009).
 Ezzell, G. A. *et al.* IMRT commissioning: Multiple institution planning and dosimetry comparisons, a report from AAPM Task
- Group 119. Med. Phys. 36, 5359–5373. https://doi.org/10.1118/1.3238104 (2009).
 37. Kimura, Y., Kadoya, N., Tomori, S., Oku, Y. & Jingu, K. Error detection using a convolutional neural network with dose difference maps in patient-specific quality assurance for volumetric modulated arc therapy. Phys. Med. 73, 57–64. https://doi.org/10.1016/j.
- ejmp.2020.03.022 (2020).
 38. Low, D. A., Harms, W. B., Mutic, S. & Purdy, J. A. A technique for the quantitative evaluation of dose distributions. *Med. Phys.* 25, 656–661. https://doi.org/10.1118/1.598248 (1998).
- Miffen, M. et al. Tolerance limits and methodologies for IMRT measurement-based verification QA: Recommendations of AAPM Task Group No. 218. Med. Phys. 45, e53–e83. https://doi.org/10.1002/mp.12810 (2018).
- Wang, Z., Bovik, A. C., Sheikh, H. R. & Simoncelli, E. P. Image quality assessment: From error visibility to structural similarity. IEEE Trans. Image Process. 13, 600–612. https://doi.org/10.1109/tip.2003.819861 (2004).
- Nioche, C. et al. LIFEx: A freeware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity. Can. Res. 78, 4786–4789. https://doi.org/10.1158/0008-5472.Can-18-0125 (2018).
- 42. Nioche, C., Orlhac, F. & Buvat, I. Texture-User Guide Local Image Features Extraction (2023).
- Sheen, H. *et al.* A radiomics nomogram for predicting transcatheter arterial chemoembolization refractoriness of hepatocellular carcinoma without extrahepatic metastasis or macrovascular invasion. *Abdom. Radiol.* https://doi.org/10.1007/s00261-020-02884-x (2021).
- Sheen, H. *et al.* Metastasis risk prediction model in osteosarcoma using metabolic imaging phenotypes: A multivariable radiomics model. *PLoS ONE* 14, 1–13. https://doi.org/10.1371/journal.pone.0225242 (2019).
- 45. Kabacoff, R. I. R in Action (Manning Publications Co., 2015).
- Siebers, J. V., Keall, P. J., Wu, Q., Williamson, J. F. & Schmidt-Ullrich, R. K. Effect of patient setup errors on simultaneously integrated boost head and neck IMRT treatment plans. *Int. J. Radiat. Oncol. Biol. Phys.* 63, 422–433. https://doi.org/10.1016/j.ijrobp. 2005.02.029 (2005).
- Zhen, H., Nelms, B. E. & Tome, W. A. Moving from gamma passing rates to patient DVH-based QA metrics in pretreatment dose QA. Med. Phys. 38, 5477–5489. https://doi.org/10.1118/1.3633904 (2011).
- Mu, G., Ludlum, E. & Xia, P. Impact of MLC leaf position errors on simple and complex IMRT plans for head and neck cancer. *Phys. Med. Biol.* 53, 77–88. https://doi.org/10.1088/0031-9155/53/1/005 (2008).
- Oliver, M., Bush, K., Zavgorodni, S., Ansbacher, W. & Beckham, W. A. Understanding the impact of RapidArc therapy delivery errors for prostate cancer. J. Appl. Clin. Med. Phys. 12, 3409. https://doi.org/10.1120/jacmp.v12i3.3409 (2011).
- Nithiyanantham, K. *et al.* Analysis of direct clinical consequences of MLC positional errors in volumetric-modulated arc therapy using 3D dosimetry system. *J. Appl. Clin. Med. Phys.* 16, 296–305. https://doi.org/10.1120/jacmp.v16i5.5515 (2015).
- Demircioglu, A. The effect of preprocessing filters on predictive performance in radiomics. *Eur. Radiol. Exp.* 6, 40. https://doi.org/ 10.1186/s41747-022-00294-w (2022).
- Moradmand, H., Aghamiri, S. M. R. & Ghaderi, R. Impact of image preprocessing methods on reproducibility of radiomic features in multimodal magnetic resonance imaging in glioblastoma. J. Appl. Clin. Med. Phys. 21, 179–190. https://doi.org/10.1002/acm2. 12795 (2020).

Author contributions

Conceptualization, H.S., H.B. S. and C.S. H.; methodology, J.S. K., C.K.; validation, C.K. and H.K.; formal analysis, H.S., H.B.S., J.K.; investigation, J.S.K; writing—original draft preparation, H.S., H.B. S; writing—review and editing, C.S. H., J.S. K.; visualization, H.S., H.B. S.; supervision, C.S. H. All authors have read and agreed to the published version of the manuscript.

Funding

This work was supported by a National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. NRF-2021R1F1A1050903, NRF-2020R1C1C1005713) and by a faculty research grant of Yonsei University College of Medicine (6-2021-0083).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/ 10.1038/s41598-023-35570-1.

Correspondence and requests for materials should be addressed to C.-S.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2023