Modulation indices for volumetric modulated arc therapy

2014 Phys. Med. Biol. 59 7315

(http://iopscience.iop.org/0031-9155/59/23/7315)

IOP Publishing | Institute of Physics and Engineering in Medicine Physics in Medicine & Biology

Phys. Med. Biol. **59** (2014) 7315–7340 doi:10.1088/0031-9155/59/23/7315

Modulation indices for volumetric modulated arc therapy

Jong Min Park[1,2,3,4], So-Yeon Park[1,2,3,5], Hyoungnyoun Kim[6], Jin Ho Kim[1,2,3], Joel Carlson[2,3,7] and Sung-Joon Ye[1,2,3,7]

[1] Department of Radiation Oncology, Seoul National University Hospital, Seoul, 110-744, Republic of Korea
[2] Institute of Radiation Medicine, Seoul National University Medical Research Center, Seoul, 110-744, Republic of Korea
[3] Biomedical Research Institute, Seoul National University College of Medicine, Seoul, 110-744, Republic of Korea
[4] Center for Convergence Research on Robotics, Advanced Institutes of Convergence Technology, Suwon, 443-270, Republic of Korea
[5] Interdiciplinary Program in Radiation Applied Life Science, Seoul National University College of Medicine, Seoul, 110-799, Republic of Korea
[6] Korea Institute of Science and Technology, Seoul, 136-791, Republic of Korea
[7] Program in Biomedical Radiation Sciences, Department of Transdisciplinary Studies, Seoul National University Graduate School of Convergence Science and Technology, Seoul, 151-742, Republic of Korea

E-mail: sye@snu.ac.kr

Received 19 June 2014, revised 29 September 2014
Accepted for publication 2 October 2014
Published 10 November 2014

Abstract

The aim of this study is to present a modulation index (MI) for volumetric modulated arc therapy (VMAT) based on the speed and acceleration analysis of modulating-parameters such as multi-leaf collimator (MLC) movements, gantry rotation and dose-rate, comprehensively. The performance of the presented MI (MI_t) was evaluated with correlation analyses to the pre-treatment quality assurance (QA) results, differences in modulating-parameters between VMAT plans versus dynamic log files, and differences in dose-volumetric parameters between VMAT plans versus reconstructed plans using dynamic log files. For comparison, the same correlation analyses were performed for the previously suggested modulation complexity score (MCS_v), leaf travel modulation complexity score (LTMCS) and MI by Li and Xing (MI Li&Xing). In the two-tailed unpaired parameter condition, p values were acquired. The Spearman's rho (r_s) values of MI_t, MCS_v, LTMCS and MI Li&Xing to the local gamma passing rate with 2%/2 mm criterion were −0.658 ($p < 0.001$), 0.186 ($p = 0.251$), 0.312 ($p = 0.05$) and −0.455 ($p = 0.003$), respectively. The values of r_s to the modulating-parameter (MLC

positions) differences were 0.917, −0.635, −0.857 and 0.795, respectively ($p < 0.001$). For dose-volumetric parameters, MI_t showed higher statistically significant correlations than the conventional MIs. The MI_t showed good performance for the evaluation of the modulation-degree of VMAT plans.

Keywords: modulation indices, volumetric modulated arc therapy, multi-leaf collimator

(Some figures may appear in colour only in the online journal)

1. Introduction

Intensity modulated radiation therapy (IMRT) and volumetric modulated arc therapy (VMAT) enable highly conformal delivery of a prescribed dose capable of controlling a tumor while minimizing complications in normal tissue through modulation of the treatment beam (Brahme *et al* 1982, Ezzell *et al* 2003, Otto 2008, Park *et al* 2012, Wu *et al* 2012, Oh *et al* 2013, Yu *et al* 2013). To control the beam intensities, IMRT modulates the movements of multi-leaf collimators (MLCs), while VMAT modulates not only the movements of MLCs, but also gantry rotation speed (GS) and dose-rate (DR) at given control points (CPs) (Ezzell *et al* 2003, Otto 2008). The degree of modulation might affect the deliverability of a treatment plan as originally intended due to mechanical limitations of the linac, as well as the accuracy limitations of dose calculation in the treatment planning system (TPS) (Giorgia *et al* 2007). Since highly modulated plans can potentially deliver unintended dose distributions to patients, various studies on the effects of the degree of modulation have been performed, although most have focused on IMRT (Mohan *et al* 2000, Llacer *et al* 2001, Webb 2001, 2003, Nicolini *et al* 2005, Giorgia *et al* 2007, McNiven *et al* 2010, McGarry *et al* 2011, Nauta *et al* 2011, Mittauer *et al* 2013, Tambasco *et al* 2013, Du *et al* 2014). In the case of VMAT, Nicolini *et al* investigated the impact of DR and GS on the delivery accuracy of VMAT by analyzing the dynamic log files registered by the linac control system (Nicolini *et al* 2011). They demonstrated improved plan quality and reduced treatment time when using a high DR, and slightly improved accuracy in delivery when using a low DR. Li and Xing introduced a modulation index (MI) as an assistive tool for station parameter optimized radiation therapy (SPORT) (Li and Xing 2013). They evaluated the degree of modulation at specific gantry angles in a VMAT plan by combining the segmental monitoring unit (MU) per gantry angle with the variation in MLC position. Masi *et al* introduced two MIs, one was a modified modulation complexity score for VMAT (MCS_v) and the other was a multiplicative combination of leaf travel (LT) and the MCS_v (LTMCS) (Masi *et al* 2013). The modulation complexity score (MCS) was originally described by McNiven *et al* to assess the modulation degree of IMRT plans (McNiven *et al* 2010). Masi *et al* modified that index to apply to VMAT by substituting CPs for segments. Both MCS_v and LTMCS were focused on the assessment of movements and shapes of MLCs. They used the results of pre-treatment quality assurance (QA) procedures to investigate the correlations between the values of MIs and plan deliverability.

Previous studies have demonstrated that it is possible to predict the deliverability of a treatment plan as intended by using the MI at the planning stage. Although the equivalent effect could be obtained through pre-treatment QA for each patient, using the MI prevents wasting of resources as it can be easily acquired at the planning stage, and hence filters out undeliverable plans due to excessive modulation immediately after planning. Various studies on MIs have been performed for IMRT, while investigation of MIs for VMAT has been limited (Mohan *et al* 2000, Llacer *et al* 2001, Webb 2001, 2003, Nicolini *et al* 2005, Giorgia *et al*

2007, McNiven *et al* 2010, Nauta *et al* 2011, Nicolini *et al* 2011, McGarry *et al* 2011, Li and Xing 2013, Masi *et al* 2013, Mittauer *et al* 2013, Tambasco *et al* 2013, Du *et al* 2014). Multi-leaf collimator movement in VMAT is different from that of dynamic IMRT since the MLC leaves of dynamic IMRT move in only one direction, while the MLC leaves of VMAT move in-and-out during delivery (Oh *et al* 2013). In addition, since VMAT delivers conformal dose distributions by modulating the MLC movements, GS and DR simultaneously, a comprehensive tool to evaluate the degree of modulation for each of the three variables, such as the MI proposed by Li and Xing (MI Li&Xing) is necessary (Li and Xing 2013). The MI Li&Xing was used to support a new kind of treatment technique suggested by the authors named SPORT, however, analysis of the MI with regards to the deliverability of VMAT was not performed.

In order to evaluate the degree of beam intensity modulation of VMAT plans, we analyzed the variations of mechanical parameters related to the modulation of beam intensity at each CP. These parameters were MLC position, GS and DR. In other words, MLC speed and acceleration, change of GS (gantry rotation acceleration) and variation of DR were investigated in this study. The presented MIs are focused on the variations of mechanical parameters under the assumption that an increase of mechanical parameter modulation results in an increase of uncertainty in delivery. In order to evaluate the modulation degree of VMAT comprehensively, we modified the MI previously suggested by Webb, which was originally introduced to evaluate the modulation degree of IMRT (Webb 2003). By adopting and modifying the MI by Webb, we tried to put variable weight on the variations of mechanical parameters. After that, we evaluated performances of the MIs presented in this study and compared the results to previously suggested MIs (MCS_v, LTMCS and MI Li&Xing) as applied to clinical VMAT plans for prostate and head and neck (H&N) cancer. The deliverability of each VMAT plan was evaluated with pre-treatment QA results, the differences in modulating parameters at each CP between treatment plans and delivery records, and the differences of dose-volumetric parameters between original plans and reconstructed plans with delivery records. The correlations between various MIs versus the pre-treatment QA results, the differences in delivery records, and the differences in dose-volumetric parameters were also investigated. The sensitivities and the specificities of each MI were evaluated with receiver operating characteristic (ROC) curves.

2. Materials and methods

2.1 . VMAT plans

Twenty VMAT plans for prostate cancer and 20 VMAT plans for H&N cancer, which were used for patient treatment in our institution (a total of 40 VMAT plans) were retrospectively selected to design and test performance of MIs for VMAT. Trilogy with millennium MLC (Varian Medical Systems, Palo Alto, CA, USA) was used for the planning and delivery of the H&N plans, while TrueBeam STx with high-definition MLC (Varian Medical Systems, Palo Alto, CA, USA) was used for the prostate plans. For fair comparison, new prostate VMAT plans using Trilogy with millennium MLC were generated for this study with the same structures and constraints as those used in the clinic. In the case of prostate VMAT plans, a primary plan was generated with a prescription dose of 50.4 Gy to the prostate and seminal vesicles with a margin of 2 cm in all directions except the posterior and inferior directions, where a margin of 1 cm was added. In addition, a boost plan was generated with a prescription dose of 30.6 Gy to the prostate alone with a margin of 0.7 cm. A daily dose of 1.8 Gy with two full arcs was used. Only primary plans were analyzed in this study. The H&N VMAT plans were generated to treat nasopharyngeal

cancer with a simultaneously integrated boost (SIB) technique, which had a total of three target volumes with a margin of 0.3 cm in all directions. Prescription doses were 67.5 Gy (2.25 Gy/fraction) for target 1, 54 Gy (1.8 Gy/fraction) for target 2 and 48 Gy (1.6 Gy/fraction) for target 3, to be delivered in 30 fractions. All VMAT plans for H&N cancer used 2 full arcs. Whole VMAT plans were generated in the Eclipse system (Varian Medical Systems, Palo Alto, CA, USA) with a 6 MV photon beam. For the optimization of the VMAT plans, the progressive resolution optimizer 3 (PRO3, version 10) algorithm was used. For dose calculation, the anisotropic analytic algorithm (AAA, version 10) was used with a calculation grid of 2.5 mm.

2.2. Calculation of time between CPs

Control points in VMAT plan Digital Imaging and Communications in Medicine (DICOM)'s are defined at equiangular positions of the gantry at intervals of 2.0341°. Thus to evaluate speeds and accelerations of the mechanical parameters between CPs, we needed to calculate the time between each CP. According to the Trilogy manufacturer specifications, the maximum GS and the maximum DR are $4.8° \, s^{-1}$ and $600 \, MU \, min^{-1}$, respectively. When delivering VMAT plans, the default GS is set to the maximum value ($4.8° \, s^{-1}$) and the delivered MU at each CP is controlled by DR variations. For specific CPs where large MUs exceeding the maximum DR ($600 \, MU \, min^{-1}$) must be delivered, the gantry rotation is slowed. Based on this information, we calculated the maximum MU able to be delivered without slowing down the rotation of the gantry at a CP, which was 4.238 MU. With the calculated maximum MU, we calculated the time between CPs as follows.

$$\mathrm{Time}_i = \begin{cases} \dfrac{2.0341\,^\circ}{4.8\,^\circ s^{-1}}, \Delta \mathrm{MU}_i \le 4.238 \\ \dfrac{\Delta \mathrm{MU}_i}{10 \ \mathrm{MU} s^{-1}}, \Delta \mathrm{MU}_i > 4.238 \end{cases}, \tag{1}$$

where Time_i is the time between the ith CP and the $(i+1)$th CP, and $\Delta \mathrm{MU}_i$ is the MU between the ith CP and the $(i+1)$th CP.

2.3. Calculation of modulation indices

2.3.1. MI evaluating the speed of MLC (MI_s). We modified the MI suggested by Webb (Webb 2003) for the evaluation of MLC speeds in VMAT plans. From each CP of the VMAT plan in DICOM format we extracted the position information of each MLC leaf. Utilizing the time information between each CP as calculated above, we calculated the speed of each individual MLC leaf ($\mathrm{MLC\ speed}_i$) between the CPs as follows.

$$\mathrm{MLC\ speed}_i = \frac{|\mathrm{MLC}_i - \mathrm{MLC}_{i+1}|}{\mathrm{Time}_i}, \tag{2}$$

where MLC_i is the position of the MLC at the ith CP.

After calculation of standard deviations of the $\mathrm{MLC\ speed}_i$ for each MLC leaf, we counted the number of cases where the $\mathrm{MLC\ speed}_i$ was greater than the acquired standard deviation multiplied by f value used in the MI by Webb, i.e., we acquired $z(f)$ as suggested by Webb. For individual MLCs, $z(f)$ was calculated as follows.

$$z_{\mathrm{MLC\_speed}}(f) = \left(\frac{1}{(N_{\mathrm{cp}} - 1)} \right) N\left(f;\ \mathrm{MLC\ speed}_i > f\sigma_{\mathrm{MLC\ speed}}\right), \tag{3}$$

where $f = 0.01, 0.02 \ldots 2$, $\sigma_{\text{MLC speed}}$ = standard deviation of the MLC speed$_i$, N_{cp} is the total number of CPs for a given VMAT plan and $N(f;\ \text{MLC speed}_i > f\sigma_{\text{MLC speed}})$ is a count of the number of changes for which MLC speed$_i > f\sigma_{\text{MLC speed}}$.

Using $z_{\text{MLC\_speed}}(f)$, the MI evaluating the speed of an individual MLC (*individual* MI$_s$) was calculated as follows.

$$\text{Individual MI}_s = \int_0^k z_{\text{MLC\_speed}}(f)\ \mathrm{d}f, \tag{4}$$

where $k = 0.2$, 0.5, 1 and 2 in this study

By the summation of each *individual* MI$_s$ for all MLC leaves, MI$_s$ for a single VMAT plan was acquired as follows.

$$\text{MI}_s = \sum_{n=1}^{120} \text{individual MI}_{s\,n}, \tag{5}$$

where n is the nth MLC leaf of the linac.

The *individual* MI$_s$ from MLC leaves which did not move during VMAT delivery did not contribute to an increase of MI$_s$ since the value of the *individual* MI$_s$ was 0 due to the MLC speed$_i$ being 0. As the degree of modulation of the VMAT plan increases, as does the value of MI$_s$. If there is no movement of the MLCs, MI$_s$ becomes 0.

2.3.2. MI evaluating both speed and acceleration of MLC (MI$_a$). In order to consider not only the speed of the MLC, but also the acceleration of the MLC, we added a term to MI$_s$. We acquired variations of MLC speeds at the ith CP, i.e., MLC accel$_i$, from the VMAT plan in DICOM format as follows.

$$\text{MLC accel}_i = \frac{\left|\text{MLC speed}_i - \text{MLC speed}_{i+1}\right|}{\text{Time}_i}. \tag{6}$$

After that, the standard deviation of the MLC accel$_i$ was calculated for each MLC and $z(f)$ was acquired as follows.

$$z_{\text{MLC\_accel}}(f) = \left(\frac{1}{(N_{\text{cp}} - 2)}\right) N\left(f;\ \textit{MLC speed}_i > f\sigma_{\text{MLC speed}} \text{ or MLC accel}_i > \alpha f\sigma_{\text{MLC accel}}\right), \tag{7}$$

where α = weighting factor for the acceleration which is $1/\textit{Time}_i$ acquired empirically, $\sigma_{\text{MLC accel}}$ = standard deviation of MLC accel$_i$ and $N(f;\ \text{MLC speed}_i > f\sigma_{\text{MLC speed}}$ or MLC accel$_i > \alpha f\sigma_{\text{MLC accel}})$ is a count of the number of changes for which MLC speed$_i > f\sigma_{\text{MLC speed}}$ or changes for which MLC accel$_i > \alpha f\sigma_{\text{MLC accel}}$.

Similar to the calculation of MI$_s$, individual MI$_a$ and MI$_a$ were calculated as follows.

$$\text{Individual MI}_a = \int_0^k z_{\text{MLC\_accel}}(f)\ \mathrm{d}f, \tag{8}$$

where $k = 0.2$, 0.5, 1 and 2 in this study.

$$\mathrm{MI}_a = \sum_{n=1}^{120} \text{individual } \mathrm{MI}_{a\,n}, \tag{9}$$

where n is the nth MLC leaf of the linac.

Similar to MI_s, the individual MI_a from MLC leaves which didn't move during VMAT delivery didn't contribute to an increase of MI_a since the value of individual MI_a was 0 due to the MLC speed_i and MLC accel_i both being 0. As the degree of modulation of the VMAT plan increases, as does the value of MI_a. If there is no movement of MLCs, the MI_a becomes 0.

2.3.3. MI evaluating the speed of MLC, acceleration of MLC, gantry rotation acceleration and DR variation comprehensively (MI_t). In order to consider MLC movement, acceleration of gantry rotation and DR variation comprehensively, MI_t was calculated. Gantry acceleration between the (i+1)th CP and the (i+2)th CP (GA_i) and the DR variation between the (i+1)th CP and the (i+2)th CP (DRV_i) were calculated as follows.

$$\mathrm{GA}_i = \left| \frac{\Delta\text{Gantry angle}_i}{\text{Time}_i} - \frac{\Delta\text{Gantry angle}_{i+1}}{\text{Time}_{i+1}} \right| \tag{10}$$

$$\mathrm{DRV}_i = |\mathrm{DR}_i - \mathrm{DR}_{i+1}|\,. \tag{11}$$

With these values, the weighting factors of the variations of GS and DR, which were W_{GA} at the (i+1)th CP ($W_{\mathrm{GA},i+1}$) and W_{MU} at the (i+1)th CP ($W_{\mathrm{MU},i+1}$) were calculated as follows.

$$W_{\mathrm{GA},i+1} = \frac{\beta}{\left\{1 + (\beta - 1)\cdot \mathrm{e}^{-\frac{\mathrm{GA}_i}{\gamma}}\right\}}, \tag{12}$$

where β is a constant which determines the range of $W_{\mathrm{GA},i+1}$ (in this study, β was set to 2, thereby $W_{\mathrm{GA},i+1}$ may take on values from 1 to 2) and γ is a constant which determines the speed of convergence to the maximum value of $W_{\mathrm{GA},i+1}$ (the value of γ was set to 2 in this study). The $W_{\mathrm{GA},i+1}$, designed in this study, is a monotonically increasing function from 1 to 2 as increasing the value of GA_i from 0 to infinity. Therefore, the $W_{\mathrm{GA},i+1}$ can be a weighting factor by multiplying by a certain value, making two times larger value than the original value when the GA_i is infinity while no changes in the original value when the GA_i is 0.

Similar to $W_{\mathrm{GA},i+1}$, $W_{\mathrm{MU},i+1}$ was calculated as follows.

$$W_{\mathrm{MU},i+1} = \frac{\beta}{\left\{1 + (\beta - 1)\cdot \mathrm{e}^{-\frac{\mathrm{DRV}_i}{\gamma}}\right\}}, \tag{13}$$

where β and γ were also set to 2.

The $z(f)$ for MI_t was calculated as follows.

$$z_{\mathrm{total}}(f) = \left(\frac{1}{(N_{\mathrm{cp}} - 2)}\right) \cdot \sum_{i=1}^{N_{\mathrm{cp}}} \left\{ N_i\left(f;\ \text{MLC speed}_i > f\sigma_{\text{MLC speed}} \text{ or MLC accel}_i > \alpha f \sigma_{\text{MLC accel}}\right) \cdot W_{\mathrm{GA},i+1} \cdot W_{\mathrm{MU},i+1} \right\}, \tag{14}$$

where $N_i(f$; MLC speed$_i > f\sigma_{\text{MLC speed}}$ or MLC accel$_i > \alpha f\sigma_{\text{MLC accel}})\cdot W_{\text{GA},i+1}\cdot W_{\text{MU},i+1}$ means if the MLC speed$_i$ is larger than the value of $f\sigma_{\text{MLC speed}}$, or the MLC accel$_i$ is larger than the value of $\alpha f\sigma_{\text{MLC accel}}$, then the value of that ith CP becomes 1 and is multiplied by $W_{\text{GA},i+1}$ and $W_{\text{MU},i+1}$. If the conditions are not met, the value of that ith CP becomes 0 and it is not counted. In other words, this is a weighted counting by $W_{\text{GA},i+1}$ and $W_{\text{MU},i+1}$. The numerators of $z_{\text{MLC\_speed}}(f)$ and $z_{\text{MLC\_accel}}(f)$ are integer values since they represent a counting of numbers. However, the numerator of $z_{\text{total}}(f)$ could have decimal places due to multiplication by the weighting factors ($W_{\text{GA},i+1}$ and $W_{\text{MU},i+1}$) which range from 1 to 2 due to the value of β of 2 in this study. If there is no modulation in GS and DR, both $W_{\text{GA},i+1}$ and $W_{\text{MU},i+1}$ become 1, this resulting $z_{\text{total}}(f)$ becoming $z_{\text{MLC\_accel}}(f)$. If there is extreme modulations in GS and DR, both $W_{\text{GA},i+1}$ and $W_{\text{MU},i+1}$ become 2, with the resulting $z_{\text{total}}(f)$ becoming 4 times larger than $z_{\text{MLC\_accel}}(f)$.

The individual MI$_t$ for each MLC and MI$_t$ was calculated as follows.

$$\text{Individual MI}_t = \int_0^k z_{\text{total}}(f)\ \mathrm{d}f, \tag{15}$$

where $k = 0.2$, 0.5, 1 and 2 in this study.

$$\text{MI}_t = \sum_{n=1}^{120} \text{individual MI}_{t\,n}\,. \tag{16}$$

As the degree of modulation of the VMAT plan increases, as does the value of MI$_a$.

2.4. Deliverability of VMAT plan

2.4.1. Pre-treatment QA with MapCHECK2. The planar dose distributions of VMAT plans were measured using a MapCHECK2 detector array (Sun Nuclear Corporation, Melbourne, FL, USA) inserted into a MapPHAN (Sun Nuclear Corporation, Melbourne, FL, USA). The reference 2D dose distribution of each VMAT plan was calculated by generating a virtual water phantom having the same dimensions as MapPHAN, and the CT number was assigned in the TPS as 455 following manufacturer recommendations. The calculation grid was 1 mm. Before measurements, the responses of each detector in MapCHECK2 were calibrated. The absolute dose of the linac was also calibrated. For gamma evaluation of the 2D dose distributions, SNC patient software (version 6.1.2, Sun Nuclear Corporation, Melbourne, FL, USA) was used. Both local and global gamma analyses were performed with gamma criteria of 3%/3 mm, 2%/2 mm, 1%/2 mm and 2%/1 mm.

2.4.2. Differences of modulating parameters between the treatment plans and the dynamic log files recorded during actual deliveries of plans. During pre-treatment QA, dynamic log files of each VMAT plan were acquired. Actual gantry angle and MUs at each CP during delivery were acquired as recorded by the linac control system. In addition, MLC positions every 50 ms were acquired from the Dynalog files. Using an in-house program written in Matlab (ver. 8.1, Mathworks Inc., Natick, MA, USA), these two kinds of log files were combined and formatted to correspond to the VMAT plan file in DICOM-RT format. The differences between each MLC position, gantry angle and MU from the VMAT plans and those from the log files were calculated at each CP and averaged for each VMAT plan.

2.4.3. Differences of dose-volumetric parameters between the treatment plans and the reconstructed plans using dynamic log files recorded during actual deliveries of plans. The dynamic log files in DICOM-RT format were imported to the TPS and the dose distributions were calculated using patient CT images. A calculation grid of 2.5 mm, which was the same as the original VMAT plan, was used. Identical structures as the original VMAT plans were used to calculate dose volume histograms (DVHs). For the target volume, the dose received by 95% of the target volume ($D_{95\%}$), $D_{5\%}$, minimum dose, maximum dose and mean dose were calculated. For prostate VMAT plans, $D_{20\%}$ of rectal wall, mean dose to rectal wall, $D_{20\%}$ of bladder, mean dose to bladder, $D_{50\%}$ of femoral head and mean dose to femoral head were acquired. For H&N VMAT plans, maximum dose to spinal cord, brain stem, each lens, optic chiasm and each optic nerve were calculated. In addition, mean dose to each parotid gland was calculated. The differences in dose-volumetric parameters between the original VMAT plans and the reconstructed VMAT plans using log files recorded during delivery were calculated.

2.5. Data analysis

2.5.1. Correlation analysis. In order to evaluate the performances of the presented MIs in this study, the results of pre-treatment QA, the differences in modulating parameters between the plans and the log files, and the differences in dose-volumetric parameters between the original plans and the reconstructed plans with log files were set as references for plan deliverability. Spearman's rank correlation coefficients (Spearman's rho, r_s) with p values were calculated with the presented MIs and the references. In order to obtain p values for Spearman's rho, we set two-tailed unpaired parameter condition. In addition, p values were computed using the exact permutation distributions because the applied data had small sample size without considering missing values. To compare the presented MIs to conventional MIs, MCS_v (Masi *et al* 2013), LTMCS (Masi *et al* 2013) and MI Li&Xing (Li and Xing 2013) were calculated and the correlations with the references were calculated.

For the correlation test with pre-treatment QA results and the modulating parameter differences, a total of 40 VMAT plans were used ($N = 40$), while a total of 20 VMAT plans (20 prostate plans and 20H&N plans) were used to test the correlation between dose-volumetric differences and the MIs ($N = 20$) since the structures of the prostate plans were different from those of the H&N plans.

2.5.2. Sensitivity and specificity of MI. To test the sensitivity and the specificity of the presented MIs, ROC curves and the area under the curve (AUC) were calculated with the passing rates of local gamma evaluation using gamma criterion of 2%/2 mm with 90% as a tolerance level.

3. Results

3.1. MI for VMAT plans

The MI_s, MI_a and MI_t corresponding to f values of 0.2, 0.5, 1 and 2 as well as MCS_v, LTMCS and MI Li&Xing are shown in table 1. As the value of f increases the presented MIs also increase. Since the H&N VMAT plans are generally more modulated than the prostate VMAT plans (Heilemann *et al* 2013), the MIs presented in this study had larger values in the H&N plans than those in the prostate plans. A similar tendency was observed in MI Li&Xing. Since the MCS_v and the LTMCS both decrease as modulation increases by their

Table 1. Modulation indices according to various *f*-factors.

	Prostate				Head and neck			
f	0.2	0.5	1	2	0.2	0.5	1	2
MI_s[a]	6.92 ± 0.65	10.01 ± 0.95	11.63 ± 1.37	12.20 ± 1.60	18.15 ± 2.70	21.09 ± 2.98	21.27 ± 2.92	21.28 ± 2.94
MI_a[b]	8.82 ± 0.94	14.98 ± 1.49	22.67 ± 2.30	27.58 ± 3.38	27.16 ± 4.29	42.03 ± 6.43	52.56 ± 7.63	53.51 ± 7.40
MI_t[c]	8.99 ± 0.93	15.28 ± 1.46	23.11 ± 2.26	28.12 ± 3.37	27.81 ± 4.36	43.03 ± 6.51	53.81 ± 7.73	54.78 ± 7.48
MCS_V[d]	0.59 ± 0.07				0.51 ± 0.07			
LTMCS[e]	0.39 ± 0.06				0.23 ± 0.06			
MI Li&Xing[f]	4250.06 ± 645.25				17132.10 ± 3054.03			

[a] MI based on the analysis of multi-leaf collimators speeds.
[b] MI based on the analysis of both speeds and accelerations of multi-leaf collimators.
[c] MI considering both speeds and accelerations of multi-leaf collimators, gantry accelerations and DR variations.
[d] Modulation complexity score.
[e] Multiplicative combination of LT and modulation complexity score.
[f] MI presented by Li and Xing.

Table 2. Gamma passing rates of VMAT plans for head and neck cancer as well as prostate cancer with various gamma criteria.

	3%/3mm[a]	2%/2 mm	2%/1 mm	1%/2 mm
Passing rate of local gamma evaluation (%)				
Prostate[b]	95.3	90.4	77.5	86.8
Head and neck[c]	94.3	87.4	73.9	82.2
p value	0.019	0.001	0.017	< 0.001
Passing rate of global gamma evaluation (%)				
Prostate	99.9	98.6	95.5	94.4
Head and neck	99.6	97.0	93.7	89.2
p value	0.076	0.002	0.036	< 0.001

[a] Gamma criterion.
[b] VMAT plans for prostate cancer.
[c] VMAT plans for head and neck cancer.

definitions (Masi *et al* 2013), the values of both MCS_v and LTMCS were higher in prostate plans than in H&N plans. The value of MI_a was larger than that of MI_s since an additional term to evaluate the acceleration of MLCs was added. Similarly, the value of MI_t was larger than that of MI_a.

3.2. Plan deliverability versus MI

3.2.1. Passing rates of pre-treatment QA versus MI. The average passing rates of both global and local gamma evaluation with gamma criteria of 3%/3 mm, 2%/2 mm, 2%/1 mm and 1%/2 mm are shown in table 2. The average passing rates of H&N plans were always lower than those of prostate plans with statistical significance with the exception of the results of global gamma evaluation with 3%/3 mm criterion.

The correlations between local gamma evaluation results and various MIs are shown in table 3. The results of 3%/3 mm criterion are not shown since the pre-treatment QA results of lowly modulated plans are not distinguishable from those of highly-modulated plans. The passing rates of local gamma evaluation with 2%/2 mm criterion versus MIs with f values of 0.2, 0.5, 1 and 2 as well as MCS_v, LTMCS and MI Li&Xing are plotted in figure 1. The values of r_s of MI_s, MI_a and MI_t to the local gamma passing rates with gamma criteria of 2%/2 mm, 2%/1 mm and 1%/2 mm are shown according to f values in figure 2.

In the case of 2%/2 mm and 1%/2 mm criterion, MI_s, MI_a and MI_t with every f value always had higher values of r_s than the conventional MIs with statistical significances, which means the MIs presented in this study were more correlated with the pre-treatment QA results than were the conventional MIs. On average, the r_s values of MI_s, MI_a and MI_t for the QA results with 2%/2 mm were −0.637, −0.648 and −0.660, respectively. Those values for the QA results with 1%/2 mm were −0.662, −0.668 and −0.675, respectively. The MI_t ($f = 0.2$) showed the highest correlation with the pre-treatment QA results with 1%/2 mm gamma criteria ($r_s > 0.7$ with $p < 0.001$). This is shown in figures 1 and 2.

In the case of 2%/1 mm criterion, All the MIs presented in this study had higher values of r_s with statistical significances than did the conventional MIs with the exception of MI_s with f values of 1 and 2, and MI_a and MI_t with f value of 2. In these cases, MI Li&Xing had higher r_s than the presented MIs. The average r_s values of MI_s, MI_a and MI_t for the QA results with 2%/1 mm were −0.507, −0.560 and −0.556, respectively.

Table 3. Correlations between local gamma passing rates with various gamma criteria and the modulation indices for VMAT plans.

		2%/2 mm		2%/1 mm		1%/2 mm	
MI	f	r_s[a]	p value	r_s	p value	r_s	p value
MI_s[b]	0.2	−0.627	< 0.001	−0.574	< 0.001	−0.669	< 0.001
	0.5	−0.636	< 0.001	−0.513	0.001	−0.657	< 0.001
	1	−0.637	< 0.001	−0.471	0.002	−0.657	< 0.001
	2	−0.646	< 0.001	−0.471	0.002	−0.664	< 0.001
MI_a[c]	0.2	−0.654	< 0.001	−0.617	< 0.001	−0.691	< 0.001
	0.5	−0.638	< 0.001	−0.585	< 0.001	−0.663	< 0.001
	1	−0.663	< 0.001	−0.561	< 0.001	−0.669	< 0.001
	2	−0.638	< 0.001	−0.476	0.002	−0.649	< 0.001
MI_t[d]	0.2	−0.665	< 0.001	−0.611	< 0.001	−0.701	< 0.001
	0.5	−0.658	< 0.001	−0.586	< 0.001	−0.675	< 0.001
	1	−0.667	< 0.001	−0.552	< 0.001	−0.669	< 0.001
	2	−0.648	< 0.001	−0.473	0.002	−0.656	< 0.001
MCS_v[e]		0.186	0.251	0.365	0.021	0.157	0.334
LTMCS[f]		0.312	0.050	0.371	0.018	0.343	0.030
MI Li&Xing[g]		−0.455	0.003	−0.490	0.001	−0.502	0.001

[a] Spearman's rho.
[b] MI based on the analysis of multi-leaf collimators speeds.
[c] MI based on the analysis of both speeds and accelerations of multi-leaf collimators.
[d] MI considering both speeds and accelerations of multi-leaf collimators, gantry accelerations and DR variations.
[e] Modulation complexity score.
[f] Multiplicative combination of LT and modulation complexity score.
[g] MI presented by Li and Xing.

In general, the value of r_s calculated from MI_t was higher than the r_s of MI_a. Similarly, the value of r_s of MI_a was higher than that of MI_s. Therefore, MI_t was the most correlated with the pre-treatment QA results, while the least correlated index with the QA results presented in this study was MI_s. However, the magnitudes of the differences were not large. For the conventional MIs, MI Li&Xing showed the best performance, while MCS_v showed the worst performance, neither performing as well as the MIs presented in this study.

3.2.2. Modulating parameter differences between treatment plans and deliveries versus MI. The average differences of the modulating parameters between treatment plans and actual deliveries of VMAT are shown in table 4. The parameter that differed the most with statistical significance ($p < 0.001$) between the prostate (relatively lowly modulated) and the H&N (relatively highly modulated) VMAT plans was MLC positional difference. The MU difference was not statistically significant ($p = 0.158$).

The correlations between the differences in modulating parameters and the various MIs are shown in table 5. The differences of MLC positions between the plan and the delivery versus MI_t with f values of 0.2, 0.5, 1 and 2, as well as MCS_v, LTMCS and MI Li&Xing are plotted in figure 3. The values of r_s of MI_s, MI_a and MI_t versus the differences of MLC positions are shown according to the f value in figure 4. The r_s values of MI_s, MI_a and MI_t with f values of 0.2 and 0.5 for the MLC position differences were each larger than 0.9 with statistical significances ($p < 0.001$ in every case), indicating strong correlations between the presented MIs and the MLC positional errors during delivery. In the cases of MIs with f values of 1 and 2, the r_s values were larger than 0.835 and 0.770, respectively, with statistical significance ($p < 0.001$

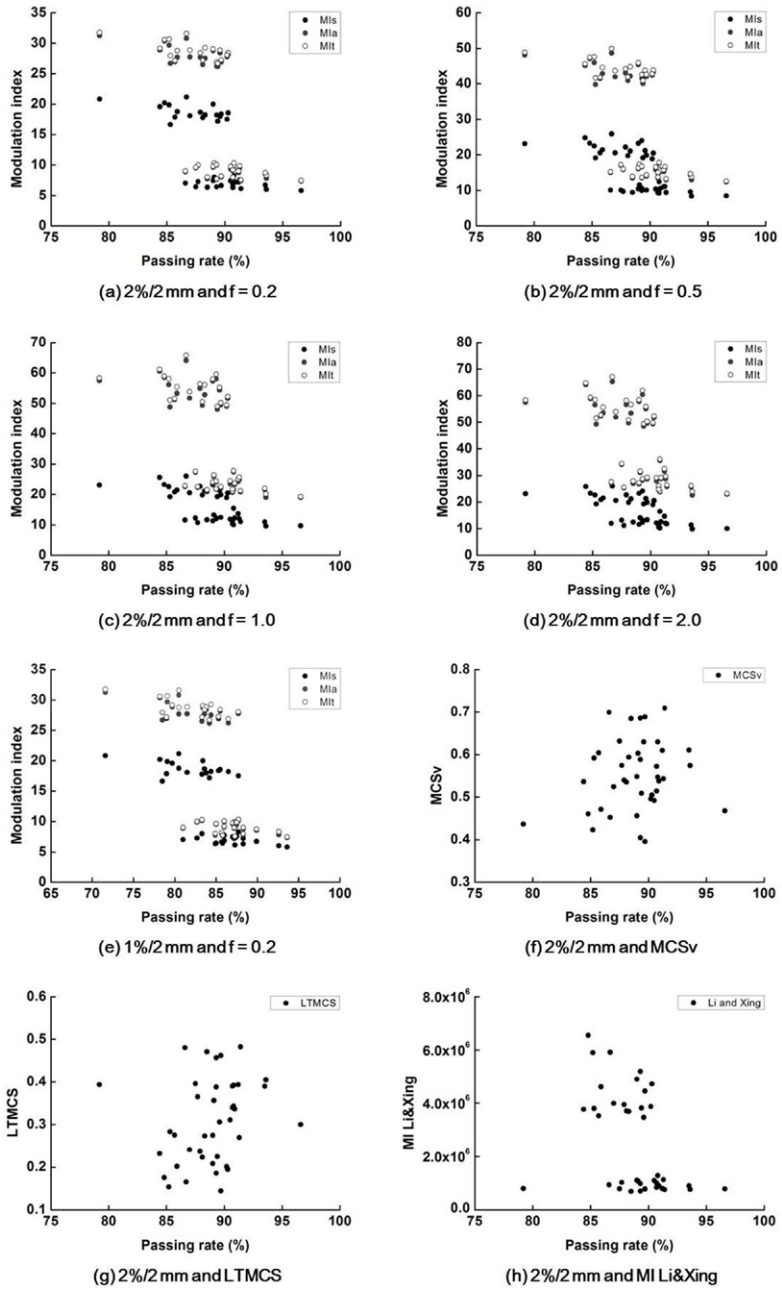

Figure 1. The passing rates of local gamma evaluation versus various modulation indices (MI) are plotted. The MI_s, MI_a and MI_t are plotted with black, gray and white circles, respectively. The MI_s, MI_a and MI_t with f value of 0.2 (*a*), 0.5 (*b*), 1 (*c*) and 2 (*d*) according to the local gamma passing rate with 2%/2 mm criterion are shown. The MI_s, MI_a and MI_t with *f* value of 0.2 according to the local gamma passing rate with 1%/2 mm criterion, which showed the highest Spearman's rho ($r_s = -0.665$ with $p < 0.001$ for MI_s, $r_s = -0.611$ with $p < 0.001$ for MI_a and $r_s = -0.701$ with $p < 0.001$ for MI_t) are also shown (*e*). The MCS_v (*f*), LTMCS (*g*) and MI Li&Xing (*h*) according to the local gamma passing rates with 2%/2 mm criterion are also shown.

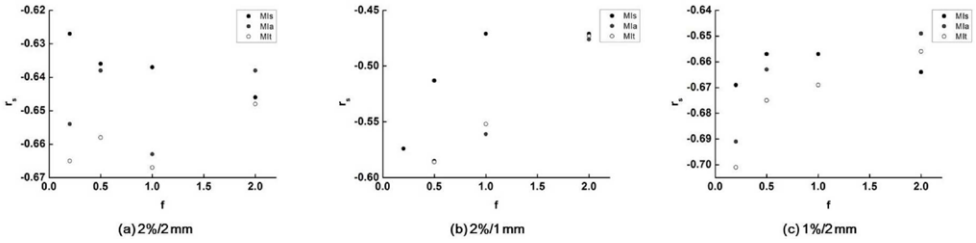

Figure 2. The Spearman's rank correlation coefficients (r_s) of MI_s, MI_a and MI_t versus the local gamma passing rates of pre-treatment quality assurance (QA) with gamma criteria of 2%/2 mm (*a*), 2%/1 mm (*b*) and 1%/2 mm (*c*) are shown according to f values of 0.2, 0.5, 1 and 2. The MI_s, MI_a and MI_t are plotted with black, gray and white circles, respectively. Every plotted r_s value showed statistical significance ($p < 0.05$).

Table 4. Average differences of modulating parameters between VMAT plans and dynamic log files recorded during actual deliveries.

	MLC[a] position (mm)	Gantry angle (°)	MU[b]
Prostate[c]	0.24 ± 0.03	0.39 ± 0.04	0.16 ± 0.02
Head and neck[d]	0.80 ± 0.04	0.37 ± 0.04	0.14 ± 0.13
p value	< 0.001	< 0.001	0.158

[a] Multi-leaf collimator.
[b] Monitoring unit.
[c] VMAT plans for prostate cancer.
[d] VMAT plans for head and neck cancer.

in every case), still showing strong correlations with the MLC positional errors during delivery. The LTMCS showed the strongest correlation with the MLC positional error during delivery among the conventional MIs ($r_s = -0.857$ with $p < 0.001$). The MCS_v and MI Li&Xing had r_s values of −0.635 with $p < 0.001$ and 0.795 with $p < 0.001$, respectively.

The MI_t was correlated with the gantry angle error in delivery, showing values of r_s higher than 0.610 with statistical significance ($p < 0.001$ in every case). The MI Li&Xing showed the strongest correlation with the gantry angle error with statistical significance ($r_s = 0.721$ and $p < 0.001$).

In the case of MU error, no correlations between the MIs and MU differences in delivery were observed.

3.2.3. Dose-volumetric differences versus MI. The average differences of dose-volumetric parameters between treatment plans and the reconstructed plans with dynamic log files recorded during delivery of VMAT for prostate and H&N cancer are shown in table 6 and table 7, respectively. The correlations between dose-volumetric parameters and the various MIs for prostate and H&N cancer are shown in table 8 and table 9, respectively.

In the case of prostate VMAT plans, the MI_s ($f = 0.5$) was correlated with the difference in $D_{95\%}$ of target ($r_s = 0.55$, $p = 0.01$), minimum dose to target ($r_s = 0.463$, $p = 0.04$) and mean dose to bladder ($r_s = 0.442$, $p = 0.05$). The MI_a and MI_t were correlated with the difference in $D_{20\%}$ of the rectal wall ($r_s = 0.496$ with $p = 0.026$ for MI_a and $r_s = 0.54$ with $p = 0.014$ for MI_t) and mean dose to rectal wall ($r_s = 0.445$ with $p = 0.049$ and $r_s = 0.479$ with $p = 0.033$). The differences in $D_{20\%}$ of the rectal wall between the original treatment plan and the reconstructed plan versus MI_t with *f* values of 0.5, as well as MCS_v, LTMCS and MI Li&Xing are plotted in

Table 5. Correlations of the average differences of modulating parameters between VMAT plans and dynamic log files with the modulation indices.

	0.2		0.5		1		2	
f	r_s[a]	*p* value	r_s	*p* value	r_s	*p* value	r_s	*p* value
Multi-leaf collimator position differences								
MI_s[b]	0.932	< 0.001	0.917	< 0.001	0.835	< 0.001	0.772	< 0.001
MI_a[c]	0.938	< 0.001	0.919	< 0.001	0.844	< 0.001	0.773	< 0.001
MI_t[d]	0.932	< 0.001	0.917	< 0.001	0.835	< 0.001	0.772	< 0.001
MCS_v[e]	−0.635	< 0.001						
LTMCS[f]	−0.857	< 0.001						
MI Li&Xing[g]	0.795	< 0.001						
Gantry angle differences								
MI_t	0.611	< 0.001	0.630	<0.001	0.638	< 0.001	0.632	< 0.001
MCS_v	−0.620	< 0.001						
LTMCS	−0.714	< 0.001						
MI Li&Xing	0.721	< 0.001						
MU differences								
MI_t	−0.136	0.404	−0.160	0.325	−0.218	0.177	−0.222	0.168
MCS_v	0.136	0.403						
LTMCS	−0.083	0.609						
MI Li&Xing	0.065	0.691						

[a] Spearman's rho.
[b] MI based on the analysis of multi-leaf collimators speeds.
[c] MI based on the analysis of both speeds and accelerations of multi-leaf collimators.
[d] MI considering both speeds and accelerations of multi-leaf collimators, gantry accelerations and DR variations.
[e] Modulation complexity score.
[f] Multiplicative combination of LT and modulation complexity score.
[g] MI presented by Li and Xing.

figure 5. The MCS_v was correlated with the difference in $D_{5\%}$ of the target ($r_s = -0.456$ with $p = 0.043$) and the maximum dose to target ($r_s = -0.548$ with $p = 0.012$). The LTMCS was correlated with the difference in the maximum dose to target ($r_s = -0.553$ with $p = 0.011$), while the MI Li&Xing was correlated with the difference in the mean dose to the rectal wall ($r_s = 0.452$ with $p = 0.045$) and the $D_{50\%}$ of the femoral head ($r_s = 0.461$ with $p = 0.041$).

In the case of H&N VMAT plans, each of MI_s, MI_a and MI_t with *f* values of 0.5 were correlated with the difference in $D_{95\%}$ of target 1, $D_{5\%}$ of target 1, mean dose to target 1, $D_{95\%}$ of target 2, $D_{5\%}$ of target 2, mean dose to target 2, $D_{95\%}$ of target 3, $D_{5\%}$ of target 3, maximum dose to the brain stem and mean dose to right and left parotid with statistical significances ($p \leq 0.05$ in every case). The differences in $D_{95\%}$ of target 3 between the original H&N VMAT plan and the reconstructed VMAT plan versus MI_t with *f* values of 0.5, as well as MCS_v, LTMCS and MI Li&Xing are plotted in figure 6. In addition, the MI_s and MI_t were correlated with the difference in mean dose to target 3, and the MI_a and MI_t were correlated with the difference in maximum dose to target 1 with statistical significances ($p \leq 0.05$). The values of statistically significant r_s values for MI_s, MI_a and MI_t with *f* values of 0.5 to the dose-volumetric parameters are shown in figure 7, with MI_t showing the highest correlation.

The MCS_v was only correlated with the difference in minimum dose to target 3, while the LTMCS was only correlated with the difference in mean dose to target 3 with statistical

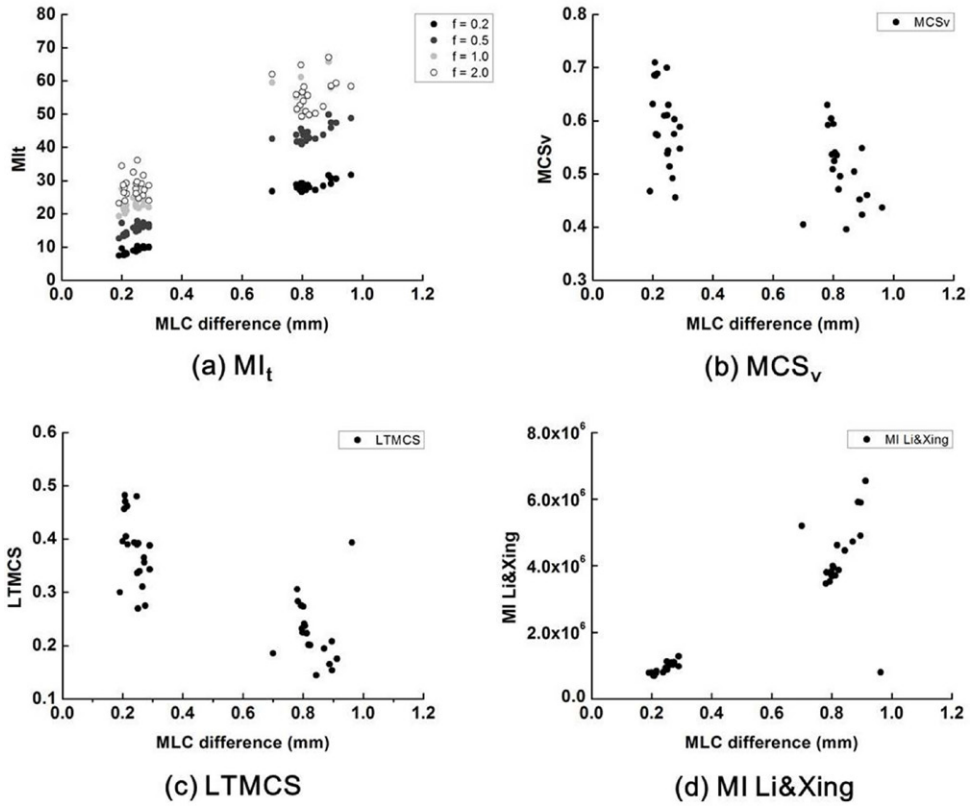

Figure 3. The differences in MLC positions between treatment plans and the dynamic log files recorded during actual deliveries versus various modulation indices (MI) are plotted. The MI$_t$ with f value of 0.2 (black circle), 0.5 (dark gray circle), 1 (gray circle) and 2 (white circle) (*a*), MCS$_v$ (*b*), LTMCS (*c*) and MI Li&Xing (*d*) are plotted against the MLC errors during delivery.

significances. The MI Li&Xing was correlated with the difference in mean dose to target 2, $D_{5\%}$ of target 3 and the mean dose to target 3 with statistical significances.

3.3. Sensitivity and specificity of MI

The ROC curves of various MIs are shown in figure 8 and the values of AUC are shown in table 10. The ROC analysis was performed with the passing rates of local gamma evaluation using 2%/2 mm criterion. A 90% passing rate was used as a tolerance level. The MI$_s$ with f value of 2 showed the best performance with an AUC value of 0.849 in terms of sensitivity and specificity. The MCS$_v$ showed the poorest performance with an AUC value of 0.527. Each of the MIs presented in this study showed larger values of AUC than the conventional MIs.

4. Discussion

In this study, we designed MIs under the assumption that an increase in the modulation of certain parameters such as MLC movement, GS and DR variations increases the uncertainty in the delivery of VMAT plans. We presented 3 different MIs, which were MI$_s$ which evaluates

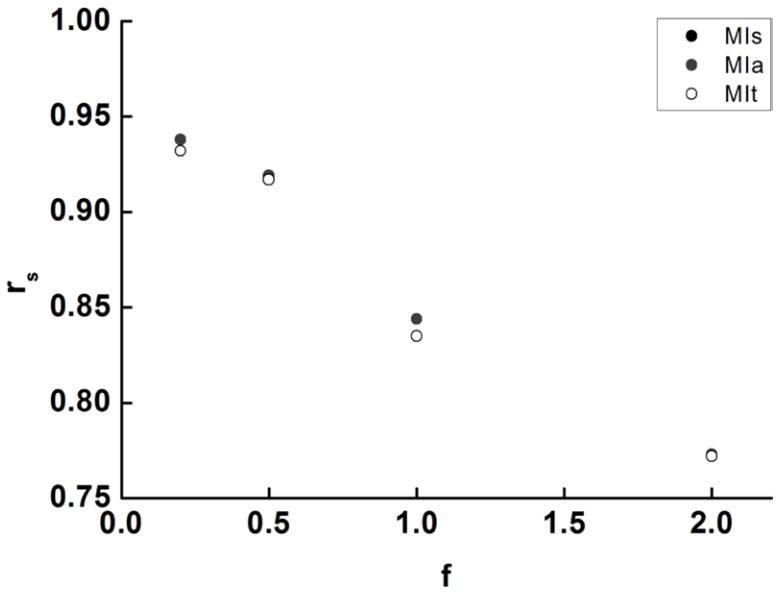

Figure 4. The Spearman's rank correlation coefficients (r_s) of MI_s, MI_a and MI_t versus the differences in MLC positions between treatment plans and the dynamic log files recorded during actual deliveries are shown according to f values of 0.2, 0.5, 1 and 2. The MI_s, MI_a and MI_t are plotted with black, gray and white circles, respectively. Every plotted r_s were statistically significant ($p < 0.05$).

Table 6. The differences in dose-volumetric parameters between VMAT plans and reconstructed VMAT plans using dynamic log files recorded during actual deliveries for prostate cancer.

Dose volumetric differences in prostate VMAT[a] plans	
Target	
$D_{95\%}$[b] (cGy)	14.0 ± 7.5
$D_{5\%}$ (cGy)	22.9 ± 9.4
Minimum dose (cGy)	12.5 ± 17.0
Maximum dose (cGy)	30.2 ± 18.4
Mean dose (cGy)	16.1 ± 7.6
Organ at risk	
Rectal wall $D_{20\%}$ (cGy)	18.5 ± 9.3
Rectal wall mean dose (cGy)	11.6 ± 5.4
Bladder $D_{20\%}$ (cGy)	2.7 ± 6.1
Bladder mean dose (cGy)	4.0 ± 2.7
Femoral head $D_{50\%}$ (cGy)	3.8 ± 4.0
Femoral head mean dose (cGy)	5.6 ± 4.0

[a] Volumetric modulated arc therapy.
[b] $D_{n\%}$ = the dose received by the $n\%$ volume of the structure volume.

MLC speed, MI_a which evaluates both speed and acceleration of MLCs, and MI_t to evaluate the speed and acceleration of MLCs, GS and DR variations, comprehensively. By adopting and modifying the methodology suggested by Webb (Webb 2003), we put variable weight

Table 7. The differences in dose-volumetric parameters between VMAT plans and reconstructed VMAT plans using dynamic log files recorded during actual deliveries for head and neck cancer.

Dose volumetric differences in head and neck VMAT[a] plans	
Target 1 (Prescription of 67.5 Gy)	
$D_{95\%}$[b] (cGy)	15.6 ± 14.0
$D_{5\%}$ (cGy)	38.4 ± 49.6
Minimum dose (cGy)	52.5 ± 53.4
Maximum dose (cGy)	37.8 ± 60.1
Mean dose (cGy)	26.6 ± 23.1
Target 2 (Prescription of 54 Gy)	
$D_{95\%}$ (cGy)	21.3 ± 10.9
$D_{5\%}$ (cGy)	32.9 ± 29.2
Minimum dose (cGy)	49.1 ± 76.4
Maximum dose (cGy)	25.7 ± 62.6
Mean dose (cGy)	26.0 ± 14.4
Target 3 (Prescription of 48 Gy)	
$D_{95\%}$ (cGy)	14.4 ± 9.0
$D_{5\%}$ (cGy)	28.9 ± 22.3
Minimum dose (cGy)	49.3 ± 56.8
Maximum dose (cGy)	26.5 ± 51.1
Mean dose (cGy)	19.9 ± 14.7
Organ at risk	
Spinal cord maximum dose (cGy)	−12.2 ± 38.9
Brain stem maximum dose (cGy)	22.0 ± 40.5
Right parotid mean dose (cGy)	42.6 ± 19.4
Left parotid mean dose (cGy)	45.4 ± 20.3
Right lens maximum dose (cGy)	11.1 ± 7.9
Left lens maximum dose (cGy)	9.0 ± 8.5
Optic chiasm maximum dose (cGy)	36.9 ± 38.0
Right optic nerve maximum dose (cGy)	25.0 ± 36.7
Left optic nerve maximum dose (cGy)	32.9 ± 27.9

[a] Volumetric modulated arc therapy.
[b] $D_{n\%}$ = the dose received by the $n\%$ volume of the structure volume.

on the variations of mechanical parameters thereby we quantified the degree of modulation of VMAT plans to evaluate the uncertainty in delivery. Since the high modulation in VMAT plans increases (1) the differences between the planned and measured dose distributions (pre-treatment QA results), (2) the differences between the modulating parameters in the treatment plan versus the dynamic log files recorded by the linac control system during actual delivery and (3) the differences between the dose-volumetric parameters from the original versus reconstructed plans from the dynamic log files, those were set as references to test the performance of the presented MIs (Agnew *et al* 2012, Heilemann *et al* 2013). The performances of the presented MIs were tested using Spearman correlation analysis between the MI values and the references. Several MIs suggested in previous studies were also calculated and their performance was compared to the presented MIs. The presented MIs showed better

Table 8. Correlations of the differences in dose-volumetric parameters between VMAT plans and reconstructed VMAT plans using dynamic log files with the modulation indices for prostate cancer.

	MI_s ($f = 0.5$)[a]		MI_a ($f = 0.5$)[b]		MI_t ($f = 0.5$)[c]		MCS_v[d]		LTMCS[e]		MI Li&Xing[f]	
	r_s[g]	p value	r_s	p value	r_s	p value	r_s	p value	r_s	p value	r_s	p value
Target												
$D_{95\%}$[h]	0.550	0.010	0.388	0.091	0.428	0.06	−0.149	0.531	−0.186	0.432	0.109	0.648
$D_{5\%}$	0.270	0.250	0.231	0.328	0.283	0.226	−0.456	0.043	−0.382	0.097	0.176	0.459
Min[i]	0.463	0.040	0.260	0.268	0.308	0.186	−0.134	0.574	−0.135	0.571	0.281	0.230
Max[j]	0.374	0.104	0.254	0.281	0.255	0.278	−0.548	0.012	−0.553	0.011	0.432	0.057
Mean[k]	0.304	0.193	0.212	0.37	0.239	0.310	−0.364	0.115	−0.336	0.147	0.156	0.511
Organs at risk												
Rectal wall $D_{20\%}$	0.406	0.076	0.496	0.026	0.54	*0.014*	−0.238	0.313	−0.280	0.232	0.350	0.130
Rectal wall mean	0.203	0.391	0.445	0.049	0.479	*0.033*	−0.232	0.324	−0.227	0.336	0.452	*0.045*
Bladder $D_{20\%}$	0.133	0.576	0.014	0.952	–0.029	0.902	−0.080	0.736	−0.121	0.610	0.072	0.761
Bladder mean	0.442	0.050	0.362	0.117	0.354	0.126	−0.141	0.554	−0.211	0.372	0.422	0.064
Femoral head $D_{50\%}$	0.114	0.631	0.227	0.335	0.246	0.295	−0.312	0.181	−0.367	0.112	0.461	0.041
Femoral head mean	0.157	0.509	0.195	0.409	0.282	0.228	−0.017	0.945	0.003	0.990	0.177	0.455

[a] MI based on the analysis of multi-leaf collimators speeds using *f*-factor of 0.5.
[b] MI based on the analysis of both speeds and accelerations of multi-leaf collimators using *f*-factor of 0.5.
[c] MI considering both speeds and accelerations of multi-leaf collimators, gantry accelerations and DR variations using *f*-factor of 0.5.
[d] Modulation complexity score.
[e] Multiplicative combination of LT and modulation complexity sc.
[f] MI presented by Li and Xing.
[g] Spearman's rho.
[h] $D_{n\%}$ = the dose received by the $n\%$ volume of the structure volume.
[i] Minimum dose.
[j] Maximum dose.
[k] Mean dose.

Table 9. Correlations of the differences in dose-volumetric parameters between VMAT plans and reconstructed VMAT plans using dynamic log files with the modulation indices for head and neck cancer.

	MI_s ($f = 0.5$)[a]		MI_a ($f = 0.5$)[b]		MI_t ($f = 0.5$)[c]		MCS_v[d]		LTMCS[e]		MI Li&Xing[f]	
	r_s[g]	p value	r_s	p value	r_s	p value	r_s	p value	r_s	p value	r_s	p value
Target 1 (Prescription of 67.5 Gy)												
$D_{95\%}$[h]	0.476	0.034	0.439	0.050	0.510	0.021	−0.288	0.218	−0.119	0.617	0.273	0.244
$D_{5\%}$	0.592	0.006	0.602	0.005	0.645	0.002	−0.279	0.233	−0.083	0.726	0.309	0.185
Min[i]	0.023	0.925	−0.012	0.960	0.034	0.887	−0.019	0.937	0.047	0.845	0.156	0.512
Max[j]	0.403	0.078	0.439	0.050	0.463	0.040	−0.201	0.396	−0.037	0.877	0.143	0.548
Mean[k]	0.562	0.010	0.543	0.013	0.580	0.007	−0.260	0.267	−0.114	0.631	0.329	0.157
Target 2 (Prescription of 54 Gy)												
$D_{95\%}$	0.529	0.016	0.472	0.035	0.548	0.012	−0.169	0.476	−0.144	0.544	0.371	0.108
$D_{5\%}$	0.644	0.002	0.605	0.005	0.689	0.001	−0.195	0.409	−0.042	0.860	0.209	0.376
Min	0.135	0.571	0.155	0.514	0.135	0.569	0.344	0.137	0.123	0.604	0.114	0.631
Max	0.117	0.624	0.108	0.652	0.122	0.609	−0.010	0.967	0.069	0.772	−0.077	0.748
Mean	0.606	0.005	0.565	0.009	0.631	0.003	−0.373	0.105	−0.322	0.167	0.517	*0.020*
Target 3 (Prescription of 48 Gy)												
$D_{95\%}$	0.595	0.006	0.599	0.005	0.674	0.001	−0.793	0.415	−0.238	0.312	0.386	0.093
$D_{5\%}$	0.554	0.011	0.505	0.023	0.563	0.010	−0.346	0.136	−0.317	0.173	0.506	*0.023*
Min	0.430	0.059	0.400	0.081	0.366	0.112	−0.479	*0.033*	−0.342	0.140	0.242	0.305
Max	0.411	0.072	0.361	0.118	0.379	0.099	−0.341	0.141	−0.392	0.087	0.419	0.066
Mean	0.477	0.033	0.422	0.064	0.482	0.031	−0.292	0.211	−0.467	*0.038*	0.590	*0.006*
Organs at risk												
Spinal cord max	0.302	0.196	0.290	0.216	0.234	0.321	−0.271	0.248	−0.070	0.769	−0.020	0.935
Brain stem max	0.500	0.025	0.519	0.019	0.512	0.021	−0.053	0.823	−0.009	0.970	0.064	0.789
Right parotid mean	0.438	0.050	0.443	0.050	0.505	0.023	−0.115	0.628	−0.135	0.571	0.359	0.120
Left parotid mean	0.470	0.036	0.447	0.048	0.444	0.050	−0.137	0.565	−0.099	0.677	0.325	0.162
Right lens max	−0.343	0.138	−0.408	0.074	−0.348	0.133	−0.113	0.634	−0.104	0.664	−0.001	0.997
Left lens max	0.052	0.827	0.026	0.914	0.116	0.625	−0.098	0.680	0.083	0.728	0.009	0.970
Optic chiasm max	−0.345	0.136	−0.391	0.089	−0.337	0.146	−0.177	0.456	−0.034	0.887	−0.017	0.942
Right optic nerve max	0.245	0.297	0.236	0.316	0.276	0.240	0.194	0.412	0.321	0.168	−0.057	0.811
Left optic nerve max	−0.341	0.141	−0.371	0.107	−0.296	0.205	0.117	0.624	0.171	0.472	−0.102	0.668

[a] MI based on the analysis of multi-leaf collimators speeds using *f*-factor of 0.5.
[b] MI based on the analysis of both speeds and accelerations of multi-leaf collimators using *f*-factor of 0.5.
[c] MI considering both speeds and accelerations of multi-leaf collimators, gantry accelerations and DR variations using f-factor of 0.5.
[d] Modulation complexity score.
[e] Multiplicative combination of LT and modulation complexity sc.
[f] MI presented by Li and Xing.
[g] Spearman's rho.
[h] $D_{n\%}$ = the dose received by the $n\%$ volume of the structure volume.
[i] Minimum dose.
[j] Maximum dose.
[k] Mean dose.

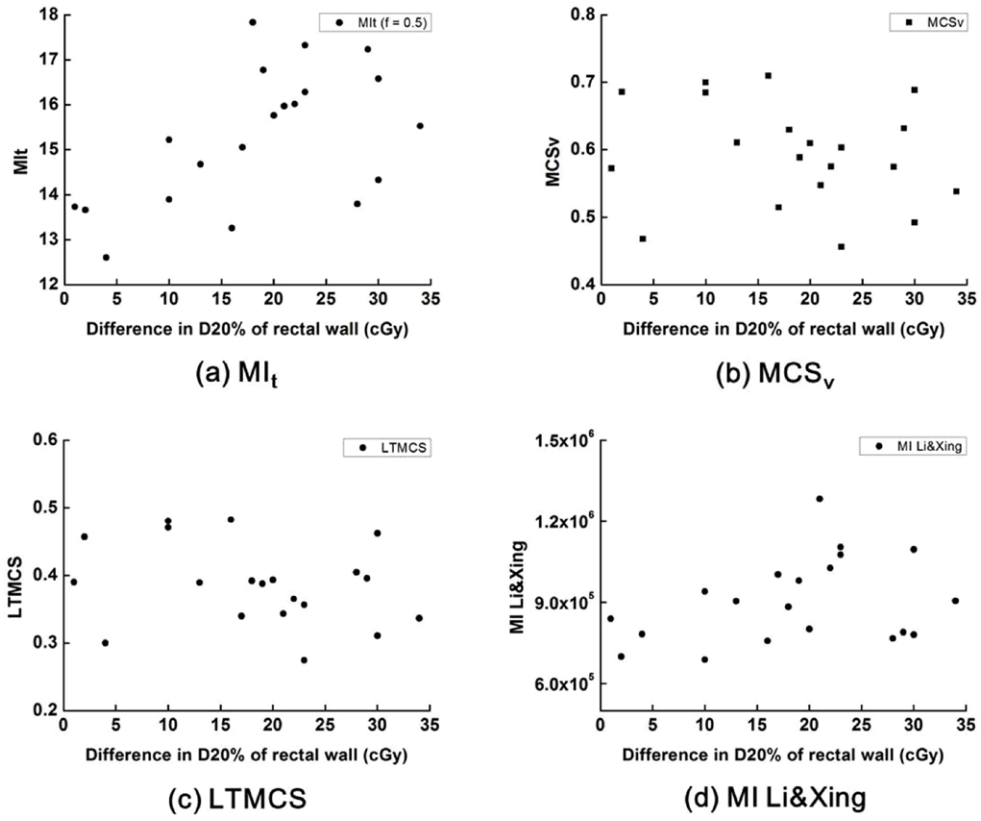

Figure 5. From the prostate VMAT plans, the differences in $D_{20\%}$ of rectal wall between the original treatment plans and the reconstructed plans with dynamic log files recorded during actual deliveries versus various modulation indices (MI_s) are plotted. The MI_t with f value of 0.5 (*a*), MCS_v (*b*), LTMCS (*c*) and MI Li&Xing (*d*) are plotted against the differences in $D_{20\%}$ of rectal wall.

performances than the conventional MIs, displaying higher values of r_s with the pre-treatment QA results, modulating parameter differences and dose-volumetric differences with statistical significances.

The presented MIs showed stronger correlations with the references than the conventional MIs. They showed better performance as predictors of poor delivery of the given treatment plans due to high modulation, as intended. This might be because the presented MIs are able to put weight on variations which are larger than a certain threshold value (equations (3), (7) and (14)) by adopting the methodology originally suggested by Webb (Webb 2003). In contrast, the conventional MIs consider every variation, even when the magnitudes are small (Li and Xing 2013, Masi *et al* 2013). Since larger variations between CPs potentially have more effect on the increase of uncertainties during delivery (Heilemann *et al* 2013), the presented MIs might show better performances than the conventional MIs by putting more weight on the larger variations. In addition, the design of the presented MIs considered the accelerations of MLC movements, gantry rotation and MU delivery. We believe that this method which evaluates the rate of changes could evaluate the uncertainty in delivery better than methods which only consider the changes of parameters in a plan. For example, in the case of MLC movements, even though the position and the speed of MLC changed minimally at certain CP,

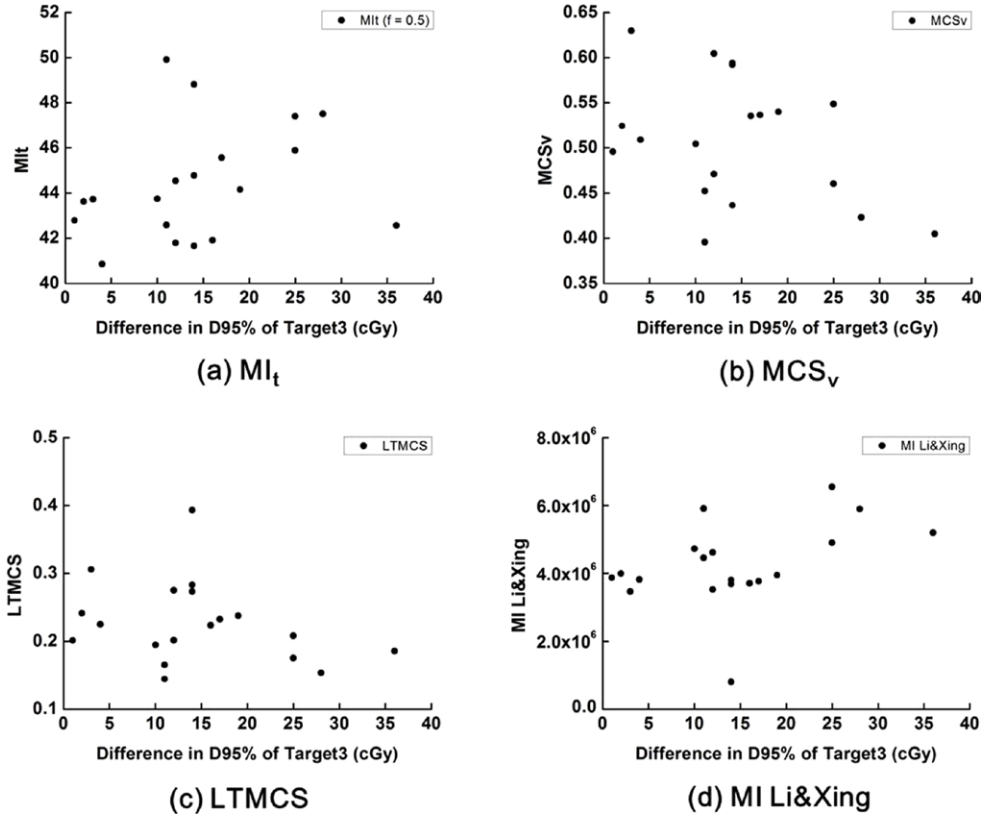

Figure 6. From the head and neck VMAT plans, the differences in $D_{95\%}$ of target 3 between the original treatment plans and the reconstructed plans with dynamic log files recorded during actual deliveries versus various modulation indices (MI_s) are shown. The MI_t with f value of 0.5 (*a*), MCS_v (*b*), LTMCS (*c*) and MI Li&Xing (*d*) are plotted against the differences in $D_{95\%}$ of target 3.

if the acceleration is large it could induce considerable uncertainty in the position of MLC at that CP during delivery. Analogous situations arise in gantry rotation and MU delivery. The MI_t presented in this study takes into account the accelerations of MLC movements, gantry rotation and MU delivery of VMAT plans comprehensively.

The MCS_v and LTMCS in this study did not show correlations as high as those found in a previous study to the pre-treatment VMAT QA results (Masi *et al* 2013). This might be due to different sample sizes and different correlation analysis methods. In the previous study performed by Masi *et al*, a total of 142 VMAT plans were analyzed using the Pearson correlation coefficient, while in this study a total of 40 VMAT plans were analyzed. Due to the smaller sample size, Spearman's rank correlation analysis was performed in this study instead of Pearson correlation analysis. In addition, the TPS, the linac system and the pre-treatment QA system used to investigate the correlation between both MCS_v and LTMCS with the pre-treatment QA results in this study were different from those used in the previous study. Even with the small sample size of 40 VMAT plans, the presented MIs showed considerable correlations to the results of pre-treatment QA with statistical significances. In the future we will perform further analysis utilizing a larger sample of VMAT plans for various tumor sites with various pre-treatment QA systems.

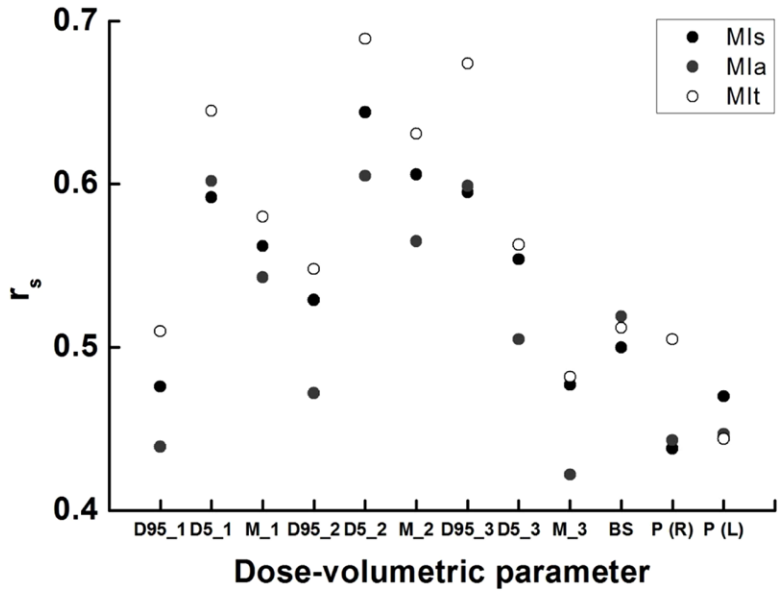

Figure 7. The Spearman's rank correlation coefficients (r_s) of MI_s, MI_a and MI_t ($f = 0.5$) versus the differences in dose-volumetric parameters between the original treatment plans and the reconstructed plans with dynamic log files recorded in the linac control system during actual deliveries for head and neck cancer are shown. The MI_s, MI_a and MI_t are plotted with black, gray and white circles, respectively. Every plotted r_s were statistically significant ($p < 0.05$). The dose received n% of the target volume for the *m*th target was expressed as Dn_m. The mean dose to the *m*th target was denoted as M_m. The BS, P (R) and P (L) means the maximum dose to brain stem, the mean dose to right parotid gland and the mean dose to left parotid gland, respectively.

In theory, the performance of MI_t should be superior in comparison with MI_s or MI_a since it contains additional terms to evaluate the GS and DR variations. By the same reasoning, MI_a should have larger values of r_s than MI_s. This was observed in the results of pre-treatment QA, even though the differences were not large. In the case of differences in MLC positions, no distinctive differences were observed since every MI presented in this study is capable of evaluating the modulation degree of MLC movements. For the differences in dose-volumetric parameters, the correlations of MI_a were generally weaker than those of MI_s, showing opposite tendency. However, the magnitudes of differences were not large. Even in this case, the MI_t showed the stronger correlations than either MI_a or MI_s. In the ROC analysis, even though the MI_s showed better performances than MI_a or MI_t, the magnitudes of the differences in AUC values among MI_s, MI_a and MI_t were not large. Therefore, the MI_t seems to be generally superior to both MI_s and MI_a. However, the magnitudes of the differences in performances between MI_t with MI_s and MI_a were not large, even though the $z_{\text{total}}(f)$ could be 4 times larger than the $z_{\text{MLC\_accel}}(f)$ or $z_{\text{MLC\_speedl}}(f)$ in the extreme case where the β value is 2. This implies that even though VMAT modulates a total of three parameters, MLC movement, GS and DR, to modulate beam intensity, the most dominant modulating parameter influencing plan deliverability of VMAT technique is the MLC movement. This is in accordance with Nicolini *et al* who demonstrated the robust compensation mechanism between GS and DR in VMAT (Nicolini *et al* 2011).

The modulating parameter differences between treatment plans and log files were strongly correlated with the presented MIs, while the pre-treatment QA results showed less, but still

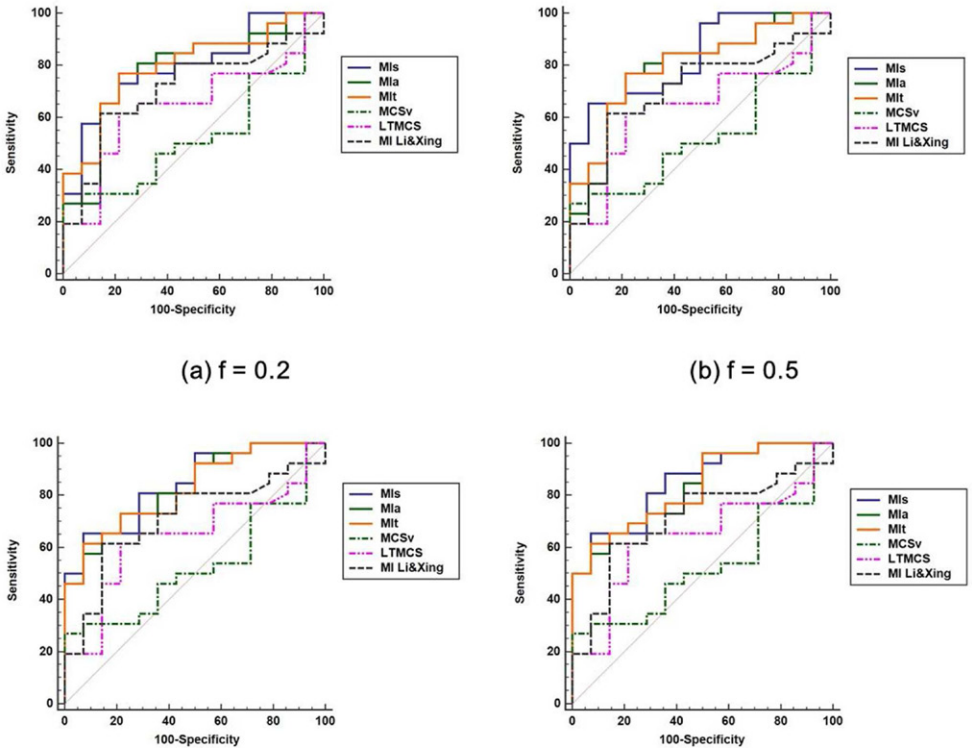

Figure 8. The ROC curves are plotted for MI_s, MI_a and MI_t with f value of 0.2 (*a*), 0.5 (*b*), 1 (*c*) and 2 (*d*) as well as MCS_v, LTMCS and MI Li&Xing with local gamma passing rates with 2%/2 mm criterion. In the case of f value of 0.2, MI_t showed the best performance while MI_s showed the best performances in the rest of the cases.

Table 10. The values of AUC from the ROC analysis for the modulation indices according to the *f* values.

f	0.2	0.5	1.0	2.0
MI_s[a]	0.794	0.832	0.843	0.849
MI_a[b]	0.783	0.791	0.830	0.830
MI_t[c]	0.797	0.799	0.824	0.830
MCS[d]		0.527		
LTMCS[e]		0.644		
Xing[f]		0.710		

[a] MI based on the analysis of multi-leaf collimators speeds.
[b] MI based on the analysis of both speeds and accelerations of multi-leaf collimators.
[c] MI considering both speeds and accelerations of multi-leaf collimators, gantry accelerations and DR variations.
[d] Modulation complexity score.
[e] Multiplicative combination of LT and modulation complexity score.
[f] MI presented by Li and Xing.

considerable, correlation. Since the magnitudes of the modulating parameter differences between the treatment plans and the dynamic log files were not large (on average, 0.52 mm difference for MLC position, 0.38° difference for gantry angle and 0.15 MU difference for

delivered MU), the effect of parameter differences on the measured dose distributions by pre-treatment QA might be minimal. In addition, this might be partially due to the limitation of 2D pre-treatment QA with gamma-index method as discussed in previous studies, as well as the insufficient resolution of the MapCHECK2 detector array (van Elmpt *et al* 2008, Wu *et al* 2012). Uncertainty of the detectors, setup of the measuring device, and the sensitivity of the detectors could cause the weaker correlation observed between the presented MIs and the pre-treatment QA results (Fredh *et al* 2013). As mentioned above, various QA systems will be investigated in further studies.

In the case of differences in dose-volumetric parameters, some parameter differences were correlated with the presented and conventional MIs, while some parameters were not. Since not every MLC movement, gantry rotation or DR variation is involved in a specific dose-volumetric parameter, the MIs could or could not have correlations with certain dose-volumetric parameters. For example, since some, but not all, of the MLCs' possible movements were involved in the sparing of dose to the brain stem in H&N VMAT plans, if the MLC leaves involved in the dose to the brain stem contributed to the change in the value of the MI, there might be a correlation. On the contrary, if the MLC leaves involved in the brain stem dose did not contribute to the change of MI and the other MLC leaves changed the value of MI, there might be no correlation. Therefore, the presented MIs and also the conventional MIs were not correlated with every dose-volumetric parameter. In general, the presented MIs were correlated with more dose-volumetric parameters than the conventional MIs, especially in H&N VMAT plans.

To determine the optimal f value for the presented MIs, values of 0.2, 0.5, 1 and 2 were tested in this study. In the case of the pre-treatment QA results, even though the f value of 0.2 showed the best performances in general, the f values of 0.5 and 1 had comparable performance. In the case of the modulating parameter differences, the f value of 0.2 showed the best performance and the f value of 0.5 showed the second best performance each with r_s larger than 0.9. In the case of the dose-volumetric parameters, after testing f values of 0.2, 0.5, 1 and 2, the MIs with f value of 0.5 showed the best performances (data are only shown for $f = 0.5$). In general, MIs with f value of 0.5 showed good performances empirically, demonstrating high correlations with the references in this study.

One limitation of this study is that we were unable to recommend a specific value of the presented MIs to objectively evaluate whether a VMAT plan is deliverable as intended or not. Even though we tested performances of the presented MIs with a total of 40 VMAT plans, those plans were all regarded as clinically acceptable and deliverable by pre-treatment QA, and had been previously used to treat patients in our institution. Since no unacceptable VMAT plans due to excessive modulation were analyzed in this study, we could not determine which value of the presented MIs could filter out clinically unacceptable plans. This will be investigated as a future work with more samples of VMAT plans as mentioned above. As a future work, clinically unacceptable VMAT plans will be acquired, and more samples will be utilized by the generation of multiple VMAT plans. The generated plans examined will have various ratios of the total MU to prescription dose per patient since VMAT plans with larger total MU may be considered more highly modulated than those with smaller total MU if the structure set and the prescription dose are identical for the same patient. In addition, for the further evaluation, the presented MIs can be compared using ANOVA according to the various pre-treatment QA conditions. Although we have addressed the similar performance of MIs in ROC analysis, the statistical significance from the large data set was not fully evaluated. By using ANOVA on a large sample size, statistically significant differences in AUC values of MIs in various conditions will be able to be analyzed as a future work, which may provide more valuable information. Another limitation of this study is that there is no consideration of the weight of small or irregular fields contained in VMAT plans, which could induce differences

between treatment plans and actual deliveries. Thus terms which include the weights of these differences will also be investigated as a future work.

5. Conclusions

The aim of this study was to design an indicator to evaluate the degree of modulation for VMAT plans focused on mechanical uncertainties during delivery to predict plan deliverability as intended. The MI_t with f value of 0.5 seems to be feasible in this study as a predictor of the plan deliverability considering the correlation with the results of pre-treatment QA, the differences in modulating parameters between treatment plans and dynamic log files recorded during actual deliveries, and the differences in dose-volumetric parameters between the original plans and the reconstructed plans with dynamic log files. The MI_t generally showed better performance in comparison with previously suggested conventional MIs, as well as MI_s and MI_a.

Acknowledgments

This work was in part supported by the National Research Foundation of Korea (NRF) grant (no. 490-20140029 and no. 490-20130047) funded by the Korea government.

References

Agnew C E, King R B, Hounsell A R and McGarry C K 2012 Implementation of phantom-less IMRT delivery verification using Varian DynaLog files and R/V output *Phys. Med. Biol.* **57** 6761–77

Brahme A, Roos J E and Lax I 1982 Solution of an integral equation encountered in rotation therapy *Phys. Med. Biol.* **27** 1221–9

Du W, Cho S H, Zhang X, Hoffman K E and Kudchadker R J 2014 Quantification of beam complexity in intensity-modulated radiation therapy treatment plans *Med. Phys.* **41** 021716

Ezzell G A, Galvin J M, Low D, Palta J R, Rosen I, Sharpe M B, Xia P, Xiao Y, Xing L and Yu C X 2003 Guidance document on delivery, treatment planning, and clinical implementation of IMRT: report of the IMRT Subcommittee of the AAPM Radiation Therapy Committee *Med. Phys.* **30** 2089–115

Fredh A, Scherman J B, Fog L S and Munck af Rosenschold P 2013 Patient QA systems for rotational radiation therapy: a comparative experimental study with intentional errors *Med. Phys.* **40** 031716

Giorgia N, Antonella F, Eugenio V, Alessandro C, Filippo A and Luca C 2007 What is an acceptably smoothed fluence? Dosimetric and delivery considerations for dynamic sliding window IMRT *Radiat. Oncol.* **2** 42

Heilemann G, Poppe B and Laub W 2013 On the sensitivity of common gamma-index evaluation methods to MLC misalignments in Rapidarc quality assurance *Med. Phys.* **40** 031702

Li R and Xing L 2013 An adaptive planning strategy for station parameter optimized radiation therapy (SPORT): segmentally boosted VMAT *Med. Phys.* **40** 050701

Llacer J, Solberg T D and Promberger C 2001 Comparative behaviour of the dynamically penalized likelihood algorithm in inverse radiation therapy planning *Phys. Med. Biol.* **46** 2637–63

Masi L, Doro R, Favuzza V, Cipressi S and Livi L 2013 Impact of plan parameters on the dosimetric accuracy of volumetric modulated arc therapy *Med. Phys.* **40** 071718

McGarry C K, Chinneck C D, O'Toole M M, O'Sullivan J M, Prise K M and Hounsell A R 2011 Assessing software upgrades, plan properties and patient geometry using intensity modulated radiation therapy (IMRT) complexity metrics *Med. Phys.* **38** 2027–34

McNiven A L, Sharpe M B and Purdie T G 2010 A new metric for assessing IMRT modulation complexity and plan deliverability *Med. Phys.* **37** 505–15

Mittauer K, Lu B, Yan G, Kahler D, Gopal A, Amdur R and Liu C 2013 A study of IMRT planning parameters on planning efficiency, delivery efficiency, and plan quality *Med. Phys.* **40** 061704

Mohan R, Arnfield M, Tong S, Wu Q and Siebers J 2000 The impact of fluctuations in intensity patterns on the number of monitor units and the quality and accuracy of intensity modulated radiotherapy *Med. Phys.* **27** 1226–37

Nauta M, Villarreal-Barajas J E and Tambasco M 2011 Fractal analysis for assessing the level of modulation of IMRT fields *Med. Phys.* **38** 5385–93

Nicolini G, Clivio A, Cozzi L, Fogliata A and Vanetti E 2011 On the impact of dose rate variation upon RapidArc implementation of volumetric modulated are therapy *Med. Phys.* **38** 264–71

Nicolini G, Fogliata A and Cozzi L 2005 IMRT with the sliding window: comparison of the static and dynamic methods. Dosimetric and spectral analysis *Radiother. Oncol.* **75** 112–9

Oh S A, Kang M K, Kim S K and Yea J W 2013 Comparison of IMRT and VMAT techniques in spine stereotactic radiosurgery with international spine radiosurgery consortium consensus guidelines *Pro. Med. Phys.* **24** 145–53

Otto K 2008 Volumetric modulated arc therapy: IMRT in a single gantry arc *Med. Phys.* **35** 310–7

Park J M, Kim J I, Heon Choi C, Chie E K, Kim I H and Ye S J 2012 Photon energy-modulated radiotherapy: Monte Carlo simulation and treatment planning study *Med. Phys.* **39** 1265–77

Tambasco M, Nygren I, Yorke-Slader E and Villarreal-Barajas J E 2013 FracMod: a computational tool for assessing IMRT field modulation *Phys. Med.* **29** 537–44

van Elmpt W, Nijsten S, Mijnheer B, Dekker A and Lambin P 2008 The next step in patient-specific QA: 3D dose verification of conformal and intensity-modulated RT based on EPID dosimetry and Monte Carlo dose calculations *Radiother. Oncol.* **86** 86–92

Webb S 2001 A simple method to control aspects of fluence modulation in IMRT planning *Phys. Med. Biol.* **46** N187–95

Webb S 2003 Use of a quantitative index of beam modulation to characterize dose conformality: illustration by a comparison of full beamlet IMRT, few-segment IMRT (fsIMRT) and conformal unmodulated radiotherapy *Phys. Med. Biol.* **48** 2051–62

Wu C *et al* 2012 On using 3D gamma-analysis for IMRT and VMAT pretreatment plan QA *Med. Phys.* **39** 3051–9

Yu M, Jang H S, Jeon D M, Cheon G S, Lee H C, Chung M J, Kim S H and Lee J H 2013 Dosimetric evaluation of tomotherapy and four-box field conformal radiotherapy in locally advanced rectal cancer *Radiat. Oncol. J.* **31** 252–9