Advances in Radiation Oncology (2020) **5**, 1032-1041

advances
in radiation oncology

www.advancesradonc.org

Scientific Article

A Risk-Adjusted Control Chart to Evaluate Intensity Modulated Radiation Therapy Plan Quality

Arkajyoti Roy, PhD,[a,*] Dan Cutright, PhD,[b] Mahesh Gopalakrishnan, MS,[c] Arthur B. Yeh, PhD,[d] and Bharat B. Mittal, MD[c]

[a]Department of Management Science and Statistics, University of Texas at San Antonio, San Antonio, Texas; [b]Department of Radiation Oncology, University of Chicago Medical Center, Chicago, Illinois; [c]Department of Radiation Oncology, Robert H. Lurie Comprehensive Cancer Center, Northwestern Memorial Hospital, Northwestern University Feinberg School of Medicine, Chicago, Illinois; and [d]Department of Applied Statistics and Operations Research, Bowling Green State University, Bowling Green, Ohio

Received 12 November 2019; revised 22 November 2019; accepted 26 November 2019

Abstract

Purpose: This study aimed to develop a quality control framework for intensity modulated radiation therapy plan evaluations that can account for variations in patient- and treatment-specific risk factors.

Methods and Materials: Patient-specific risk factors, such as a patient's anatomy and tumor dose requirements, affect organs-at-risk (OARs) dose-volume histograms (DVHs), which in turn affects plan quality and can potentially cause adverse effects. Treatment-specific risk factors, such as the use of chemotherapy and surgery, are clinically relevant when evaluating radiation therapy planning criteria. A risk-adjusted control chart was developed to identify unusual plan quality after accounting for patient- and treatment-specific risk factors. In this proof of concept, 6 OAR DVH points and average monitor units serve as proxies for plan quality. Eighteen risk factors are considered for modeling quality: planning target volume (PTV) and OAR cross-sectional areas; volumes, spreads, and surface areas; minimum and centroid distances between OARs and the PTV; 6 PTV DVH points; use of chemotherapy; and surgery. A total of 69 head and neck cases were used to demonstrate the application of risk-adjusted control charts, and the results were compared with the application of conventional control charts.

Results: The risk-adjusted control chart remains robust to interpatient variations in the studied risk factors, unlike the conventional control chart. For the brainstem, the conventional chart signaled 4 patients with unusual (out-of-control) doses to 2% brainstem volume. However, the adjusted chart did not signal any plans after accounting for their risk factors. For the spinal cord doses to 2% brainstem volume, the conventional chart signaled 2 patients, and the adjusted chart signaled a separate patient after accounting for their risk factors. Similar adjustments were observed for the other DVH points when evaluating brainstem, spinal cord, ipsilateral parotid, and average monitor units. The adjustments can be directly attributed to the patient- and treatment-specific risk factors.

Conclusions: A risk-adjusted control chart was developed to evaluate plan quality, which is robust to variations in patient- and treatment-specific parameters.

Sources of support: This work had no specific funding.

Disclosures: The data used in this study are from de-identifiable Digital Imaging and Communications in Medicine files.

* Corresponding author: Arkajyoti Roy, PhD; E-mail: arkajyoti.roy@utsa.edu

https://doi.org/10.1016/j.adro.2019.11.006

Introduction

Typically, treatment plan evaluation in intensity modulated radiation therapy (IMRT) is performed by comparing the patient's dose distribution and volume coverage to established thresholds. Often, the thresholds need to be relaxed owing to patient- and treatment-specific risk factors, such as the complexity of a patient's anatomy, tumor dose requirements, use of chemotherapy, or surgery. This process involves subjective decisions, leading to large variations in IMRT planning criteria across patients and making the planning process irreproducible and plan quality incomparable among patients.[1]

Quality control of the treatment planning process can be captured using a control chart framework, as discussed by Chung et al[2] and Pawlicki et al.[3,4] Most commonly, control charts have been used in quality assurance of dosimetry from the treatment planning system (TPS). Individual, moving-range, standard deviation (SD), and exponentially weighted moving average (EWMA) charts were used to detect unusual trends in dose deviations between the TPS and delivered dose.[5,6] Alternatively, Nordström et al[7] found control charts to be effective for the continuous verification of monitor units and derived control limits that agree with international guidelines. However, the use of control charts for treatment plan evaluation has been limited due to planning process variability for each patient, including patient- and treatment-specific risk factors.

Alternatively, plan evaluations based on patient anatomy have been explored under the knowledge-based planning (KBP) framework. Regression and machine-learning models were used to predict organ-at-risk (OAR) dose based on structure volumes, the planning target volume (PTV) and OAR volume overlap, distance-to-target histograms (DTH), and minimum distance between OAR and PTV.[8-10] In addition, a principal component analysis was used to correlate OAR sparing to anatomic features, such as the median distance between OAR and PTV, portion of OAR volume within an specific distance, OAR and PTV overlap, portion of OAR volume outside the primary treatment field, and DTH.[11,12] Additionally, overlap volume histograms (OVHs) have been used to capture the spatial configuration of an OAR with respect to a target and suggest possible improvements in OAR dose-volume histograms (DVHs), which were confirmed by replanning.[13] For further examples of plan evaluation under the KBP framework, refer to Ge and Wu.[14]

In practice, RapidPlan provides a KBP implementation to predict DVHs with upper and lower bounds at ±1SD.[15] These bounds account for some routine planning process variability, but using ±1SD under normally distributed data substantially increases the probability of a false alarm (type-I error) compared with using ±3SD, as typically used in control chart methodologies.[16] In other words, using ±1SD leads to many more false-positives, which require unnecessary reviews and clinical resources. Furthermore, KBP cannot identify trends in planning criteria because time is not considered. To address these limitations, we propose merging modeling techniques from KBP with control charts using ±3SD limits and plotting plans in chronological order.

In this proof of concept, we developed a risk-adjusted control chart that can monitor plan quality after accounting for patient- and treatment-specific risk factors. The proposed control chart employs a risk-adjustment model and a variety of risk factors that are discussed in the following. In addition, we provide details of the patient set used for demonstration. Finally, the risk-adjusted control chart is demonstrated in comparison with conventional control charts.

Methods and Materials

Risk-adjusted control charts were developed for 6 OAR DVH points and average monitor units that serve as proxies for plan quality. The 6 DVH points used are D_2, D_{20}, D_{40}, D_{60}, D_{80}, and D_{98}, where D_x represents the amount of the dose delivered to x% of the structure volume. Alternatively, the dose to the absolute volume can be used in this framework. Average monitor units were computed by dividing the total number of monitor units by the number of beams in the plans. To develop the risk-adjusted control charts, we first modeled plan quality using the patient- and treatment-specific risk factors. Then, we used the residuals from the risk adjustment model to develop the risk-adjusted charts.

Risk adjustment model

Patient-specific risk factors, such as complex anatomic geometry or high-dose prescriptions, can have an impact on OAR D_x or monitor units.[7-12] Treatment-specific risk factors, such as concurrent chemotherapy or surgery, are clinically relevant when evaluating planning criteria. We

model for the effect of such risk factors on OAR $D_{x,\,i}$ for patient i using multiple regression via

$$D_{x,i} = \beta_0 + \sum_{j=1}^{J} \beta_j z_{i,j} + \varepsilon_{x,i}, \quad (1)$$

where β_0 is the intercept, β_j is the coefficient associated with the j^{th} risk factor $z_{i,\,j}$, and residuals ε are assumed to follow a normal distribution $N(0,\sigma^2)$ with constant variance. To achieve the best fit model, backward elimination was employed for variable selection. For patient i, the residual $\varepsilon_{x,i}=D_{x,i}-\widehat{D}_{x,i}$ is the difference between the observed $D_{x,i}$ (which ignores the risk factors) and the predicted $\widehat{D}_{x,i}$ (which accounts for the risk factors). Therefore, the residual $\varepsilon_{x,i}$ represents an adjusted $D_{x,i}$ that effectively is free from the influence of risk factors $z_{i,j}$ Then, the residuals are used within a control chart framework to evaluate and monitor plan quality. Similarly, a risk adjustment model is developed for average monitor units, and the residuals are computed based on the predicted and actual monitor units.

Risk-adjusted control chart

Conventional quality control charts plot the observations, in this case OAR D_x or average monitor units. Residuals can be used as the plotting statistic if observations change depending on risk factors, and by plotting residuals that are statistically independent of risk factors and plan quality proxy, the effect of the risk factors is removed.[17]

Individual control charts can be used to evaluate and monitor individual observations, such as patient i's residual $\varepsilon_{x,i}$ (Eq. 1). The residual chart (hereinafter the risk-adjusted control chart) consists of a center line (CL) estimated as the mean of the residuals, $\overline{\varepsilon_x}$, and lower and upper control limits (LCL and UCL). The control limits are ± 3SD from the CL, yielding a very low probability of misclassifying a typical plan as unusual. The SD is estimated using the average moving range $\overline{MR}$ of all consecutive residual pairs, where the moving range for patient i, $MR_i=|\varepsilon_{x,i+1}-\varepsilon_{x,i}|$ is the absolute difference in residuals between patients i and $i+1$. $\overline{MR}$ is standardized by the constant $d_2=1.128$ to estimate SD (Eqs. 2 and 3). The constant d_2 is based on the number of consecutive points used in the MR_i computation (ie, 2 in this study, $\varepsilon_{x,i}$ and $\varepsilon_{x,i+1}$).[18] Therefore, the LCL and UCL are

$$LCL = \overline{\varepsilon_x} - 3\frac{\overline{MR}}{d_2}, \quad (2)$$

and

$$UCL = \overline{\varepsilon_x} + 3\frac{\overline{MR}}{d_2}. \quad (3)$$

The risk-adjusted control chart can identify residuals outside of the control limits (ie, out-of-control [OC]). This indicates an unusual plan quality, signaling the need for further review because the unusual observation likely cannot be attributed to the risk factors.

Risk factors

When modeling plan quality, various risk factors can be used, including anatomy, treatment history, TPS parameters, physician's experience, and institutional protocols. Here, we considered tumor dose requirements, use of chemotherapy, surgery, and the patient's anatomy parametrized through 4 geometry- and 2 proximity-based parameters. As described in Deasy et al,[19] we computed cross-sectional areas, volumes, spreads (maximum distance in x, y, or z dimensions), and surface areas of the PTV and OARs to capture the anatomic geometry. We computed the nonnegative minimum distance between the PTV and OARs, as well as the distance between PTV and OAR centroids to capture the proximity of structures.[19] When the OAR overlaps the PTV, the OAR-PTV overlap volume replaces the null minimum distance. For tumor dose requirements, 6 PTV DVH points D_2, D_{20}, D_{40}, D_{60}, D_{80}, and D_{98} are used. Therefore, a risk adjustment model for a plan quality proxy consists of $j=1,\ldots,18$ risk factors: 6 PTV DVH points, 4 PTV and 4 OAR geometric measures, 2 proximity measures, use of chemotherapy, and surgery. When modeling the average monitor units, all OAR geometry- and proximity-based risk factors were employed in the original model.

Clinical data

A total of 69 head and neck cancer cases were used for this study. Brainstem and spinal cord were identified as the primary OARs to evaluate. Descriptive statistics about the brainstem and spinal cord D_x are provided in Table 1. Ipsilateral parotid was also identified as an organ of importance (ipsilateral parotid descriptive statistics are described in the supplemental material (Table E1 and E2) available online at https://doi.org/10.1016/j.adro.2019.11.006, accompanying the statistics describing average monitor units). To parametrize tumor dose, geometry, and proximity consistently, the largest PTV was considered the primary target for each patient. The selection of the largest PTV was deemed reasonable because PTVs were treated sequentially for each patient; therefore, the largest PTV encompassed all other PTVs in the patient's treatment course.

All patients underwent a computed tomography scan (3-mm slice width) for treatment planning. All target and OAR volumes were delineated and approved by the

Table 1 Descriptive statistics of brainstem and cord D_x

	Brainstem						Spinal cord					
	D_2	D_{20}	D_{40}	D_{60}	D_{80}	D_{98}	D_2	D_{20}	D_{40}	D_{60}	D_{80}	D_{98}
Minimum	5.50	2.70	1.70	1.30	0.90	0.70	25.10	6.90	4.70	3.90	0.30	0.10
Maximum	65.10	58.70	53.90	48.70	42.30	31.30	57.10	54.10	51.90	49.70	41.70	31.10
Mean	42.84	35.18	26.16	16.51	9.18	5.64	43.36	39.37	36.44	32.39	18.81	3.62
Median	44.70	37.10	26.90	14.70	5.70	3.90	42.90	39.50	36.70	32.10	19.50	1.70
Standard deviation	9.22	10.47	11.34	11.16	8.49	6.09	4.48	5.62	5.62	5.97	11.60	5.22

Abbreviations: D_x = doses to x% brainstem volume.

same physician to reduce variability in contouring and minimize subjective preferences in planning criteria.[20] Outlier detection was performed on the anatomic data to identify inconsistencies in contouring, and resulted in the redelineation of 3 plans. Treatment planning was performed using Pinnacle[3] (Philips) versions 6.2b to 9.10. Treatments were delivered with step-and-shoot intensity modulation. Radiation doses were prescribed for sequential delivery using a standard fractionation scheme of 1.8 to 2.0 Gy or 1.5 Gy twice daily. The low-risk target volume received 45 to 50 Gy while the intermediate-to-high risk target volume received 56 to 66 Gy. The gross tumor volume typically received 70 to 72 Gy.

During the planning process, every attempt was made to ensure that 100% of the prescribed dose covered at least 95% of the target volume with a maximum dose <110% of the prescribed dose. For OARs and normal tissue, a combination of maximum dose and maximum DVH objectives were used for optimization. When serial OARs were in close proximity to treatment volumes, priority was given to these OARs while still maintaining adequate coverage for target volumes. A composite DVH was developed from the sequential deliveries for evaluation. The composite plan data were exported to a computational environment for radiation therapy research for data analysis via Digital Imaging and Communications in Medicine (DICOM) file export from Pinnacle[3,19]

Results

First, we provide the descriptive statistics of all risk factors included in the modeling. Then, we focus on the risk adjustment model and control charts for brainstem and spinal cord D_2 because this reflects the maximum dose criteria commonly used to evaluate brainstem and spinal cord doses for head and neck cases.[21] The regression outputs, diagnostic plots, and control charts for D_{20}, D_{40}, D_{60}, D_{80}, and D_{98} on the brainstem and spinal cord accompanied the 6 DVH points for the ipsilateral parotid and average monitor units in the supplemental material (available online at https://doi.org/10.1016/j.adro.2019.11.006).

Risk factors

The descriptive statistics on the PTV DVH points for D_2, D_{20}, D_{40}, D_{60}, D_{80}, and D_{98} for the 69 head and neck cases are summarized in Table 2. One patient received a low dose to the PTV, leading to a minimum $D_{40}=8.2$, $D_{60}=3.6$, $D_{80}=2.6$, and $D_{98}=0.1$. This was unusual compared with the remaining plans, with the second lowest $D_{40}=58.2$, $D_{60}=51.1$, $D_{80}=42.6$, and $D_{98}=1.6$.

Statistics on the PTV, brainstem, and spinal cord geometries are summarized in Table 3. The mean cross-sectional area, volume, and surface area for the PTV were larger than the brainstem and spinal cord. However, the spinal cord spread was observed to be larger than the brainstem and the PTV.

Statistics on proximity measures, such as minimum and centroid distances between the OARs and PTV are summarized in Table 4. One patient had a PTV that overlapped with the brainstem, leading to a no-minimum distance between the structures. All remaining patients had positive minimum distances between the PTV and OARs. Furthermore, the average distance between the brainstem and PTV centroids was larger than the average distance between the spinal cord and PTV centroids. Treatment history was incorporated by accounting for concurrent chemotherapy (66 patients) and surgery (39 patients). These geometry-, proximity-, PTV dose-, and treatment-based risk factors are incorporated into the risk adjustment model.

Table 2 Descriptive statistics of PTV D_x

	D_2	D_{20}	D_{40}	D_{60}	D_{80}	D_{98}
Minimum	61.5	37.0	8.2	3.6	2.6	0.1
Maximum	83.1	78.6	76.7	72.8	67.8	60.2
Mean	76.0	72.2	68.5	62.3	53.9	45.6
Median	76.4	73.3	70.1	62.8	54.5	47.2
Standard deviation	3.6	5.7	8.5	8.6	8.2	8.9

Abbreviations: D_x = doses to x% brainstem volume.

Table 3 Minimum, mean, median, maximum, and standard deviation of the PTV, brainstem, and spinal cord geometries

Structure	Cross-sectional area (cm^2)					Volume (cm^3)				
	Min	Mean	Median	Max	SD	Min	Mean	Median	Max	SD
PTV	27.0	68.7	67.0	145.1	20.5	289.9	1317.7	1254.5	2451.5	448.0
Brainstem	1.7	4.2	4.1	10.5	1.5	9.4	27.4	26.4	47.3	6.7
Spinal cord	0.5	1.0	1.0	1.8	0.3	10.8	21.3	21.0	34.4	4.8
Structure	Spread (cm)					Surface area (cm^2)				
	Min	Mean	Median	Max	SD	Min	Mean	Median	Max	SD
PTV	10.2	20.5	20.7	28.6	3.3	239.2	969.1	974.4	1477.0	229.1
Brainstem	3.3	5.8	6.0	7.5	0.8	24.3	57.4	56.7	84.3	10.1
Spinal Cord	14.4	20.1	20.1	24.3	2.1	65.6	126.0	126.8	186.8	24.5

Abbreviations: max = maximum; min = minimum; PTV = planning target volume; SD = standard deviation.

Risk adjustment model

The regression output from the risk adjustment model for the brainstem and spinal cord D_2 are reported in Tables 5 and 6, respectively, with accompanying diagnostic plots shown in Figure 1. Although all 18 risk factors were included in every OAR model, the final model with the best fit risk factors using the Akaike information criterion is reported in the regression output in Tables 5 and 6. The Akaike information criterion is most commonly used for model selection, and measures the relative quality of the regression model when risk factors are removed iteratively to improve model quality.[22]

When modeling brainstem D_2, model fit and significance are reported by $R^2 = 81.6\%$ and $P = 2.2e\text{-}16$, respectively, as shown in Table 5. For spinal cord D_2, $R^2 = 49.38\%$ and $P = 7.47e\text{-}06$ (Table 6). Both best fit models produced significant results; however, the fit of the brainstem model was substantially higher than the spinal cord model.

For both risk adjustment models, the assumptions of linearity, normality, and homoscedasticity appear valid. The residual versus fitted values plot in Figure 1A shows a random scattering of residuals around a mean of 0 and within a horizontal band, demonstrating linearity in the relationships between x and y, and homoscedasticity in the residuals. The normal Q–Q plot shows the residuals lying approximately on the $y=x$ line in Figure 1B, demonstrating the normality of the residuals. Similar observations are made in Figures 1C and D to reflect valid assumptions for spinal cord D_2

The residuals from the regression models are used as the charting variable for the risk-adjusted control charts as described in the following section.

Risk-adjusted control chart

Conventional individual control charts have been used for the quality control of planning parameters in radiation therapy. Therefore, we compared the results from the proposed risk-adjusted control charts with those of the conventional individual control charts for the brainstem and spinal cord D_2 Figure 2A demonstrates the conventional chart plotting brainstem D_2, with an $UCL=63.5$Gy and $LCL=22.2$ Gy. Patient 4 was identified as out OC based on a high $D_2=65.1$ Gy compared with the remaining plans. Conversely, patients 42, 51, and 52 were identified as OC based on their low brainstem $D_2=20.3$Gy, 6.1Gy, and 5.5Gy, respectively.

Figure 2B displays the conventional chart for spinal cord D_2, with $UCL=54.2$Gy and $LCL=32.5$Gy. Patient 1 was identified as OC based on their spinal cord $D_2=57.1$Gy that exceeds the UCL. Patient 43 was identified as OC based on their spinal cord $D_2=25.1$Gy that is unusually low compared with the remaining patients.

For our proposed method, Figure 2C depicts the risk-adjusted individual chart plotting the residuals from the brainstem D_2 model, with $UCL=11.6$ Gy and $LCL=-11.6$Gy. Unlike the conventional chart, no patients had an OC brainstem D_2 after accounting for the best fit risk factors. Figure 2D illustrates the risk-adjusted control chart for spinal cord D_2 with $UCL=9.0$ Gy and

Table 4 Descriptive statistics of proximity between PTV and organs at risk

	Brainstem – PTV		Spinal cord – PTV	
	Minimum distance (cm)	Distance between centroids (cm)	Minimum distance (cm)	Distance between centroids (cm)
Minimum	0.0	6.5	0.5	3.9
Maximum	2.6	12.5	1.7	14.0
Mean	0.8	10.1	0.9	6.0
Median	0.7	10.2	0.9	5.8
Standard deviation	0.4	1.2	0.2	1.4

Abbreviation: PTV = planning target volume.

Table 5 Regression output for modeling brainstem D_2

Coefficients	Estimate	Standard error	*t* statistic	*P*-value
Intercept	74.7313	8.1357	9.186	6.56E-13
PTV cross-sectional area	−0.1671	0.0412	−4.057	0.00015
PTV	0.0119	0.0024	4.953	6.66E-06
PTV surface area	−0.0040	0.0026	−1.556	0.125181
Minimum distance brainstem − PTV	−13.4698	1.5823	−8.513	8.52E-12
Centroid distance brainstem − PTV	−2.9535	0.7141	−4.136	0.000115
Brainstem volume	−0.2113	0.1070	−1.975	0.052981
Brainstem spread	2.4817	0.8362	2.968	0.004352
PTV D_{40}	−0.3121	0.2250	−1.387	0.170833
PTV D_{60}	0.8202	0.3271	2.507	0.014983
PTV D_{80}	−0.5501	0.1853	−2.969	0.004341
Multiple R^2	81.58%		*f* statistic	25.69
Adjusted R^2	78.40%		*P*-value	2.2e-16
Residual standard error	4.284			

Abbreviations: D_x = doses to x% brainstem volume; PTV = planning target volume.

$LCL = -9.0$ Gy, where only patient 19 was identified as an unusual plan with a low spinal cord $D_2 = 36.9$ Gy.

For the remaining brainstem D_X, the risk-adjusted control chart identified patients 4 and 58 with unusually high D_{80} and D_{98}, and patient 61 was identified with unusually low D_{20} (Fig E4; available online at https://doi.org/10.1016/j.adro.2019.11.006). For the remaining spinal cord D_x, the plans for patients 1 and 3 produced an unusually high spinal cord D_{98} (Fig E5; available online at https://doi.org/10.1016/j.adro.2019.11.006).

For the ipsilateral parotid, patient 1 was OC owing to a low D_2 and a low D_{20}, patient 51 was OC due to a low D_{40}, and patients 52 and 54 were OC due to a low D_{60} (Fig E6; available online at https://doi.org/10.1016/j.adro.2019.11.006). When evaluating average monitor units, 2 plans were identified as OC in the conventional chart, with patients 31 and 61 having an unusually high number of average monitor units. Patient 31 remained OC in the risk-adjusted chart, and patient 61 was adjusted to be in control (Fig E7; available online at https://doi.org/10.1016/j.adro.2019.11.006).

Discussion

Many patients were identified as OC in the conventional charts across the DVH points for all 3 OARs and average monitor units. This was expected owing to the large variability in the DVHs and monitor units across patients. The risk-adjusted control chart typically reduced the number of OC plans by attributing some of the variability between patients to their respective risk factors. For example, compared with the entire patient set, the minimum distances between the PTVs and brainstems for patients 51 and 52 were >3.4 SD above the mean. This contributes to the unusually low brainstem D_2. Similarly,

Table 6 Regression output for modeling spinal cord D_2

Coefficients	Estimate	Standard error	*t* statistic	*P*-value
Intercept	52.9373	11.0490	4.791	1.19E-05
PTV cross-sectional area	−0.0697	0.0311	−2.241	0.0289
PTV	0.0073	0.0023	3.238	0.00199
PTV spread	−0.4188	0.2779	−1.507	0.13728
PTV surface area	−0.0034	0.0021	−1.638	0.10689
Minimum distance cord − PTV	−14.2842	2.3202	−6.157	7.51E-08
Cord volume	−0.1930	0.0959	−2.013	0.04881
PTV D_2	0.3003	0.1445	2.078	0.04213
PTV D_{40}	−0.4475	0.1873	−2.389	0.02019
PTV D_{60}	0.8087	0.2585	3.128	0.00275
PTV D_{80}	−0.5125	0.1526	−3.358	0.00139
Multiple R^2	49.38%		*f* statistic	5.657
Adjusted R^2	40.65%		*P*-value	7.47e-06
Residual standard error	3.452			

Abbreviations: D_x = doses to x% brainstem volume; PTV = planning target volume.

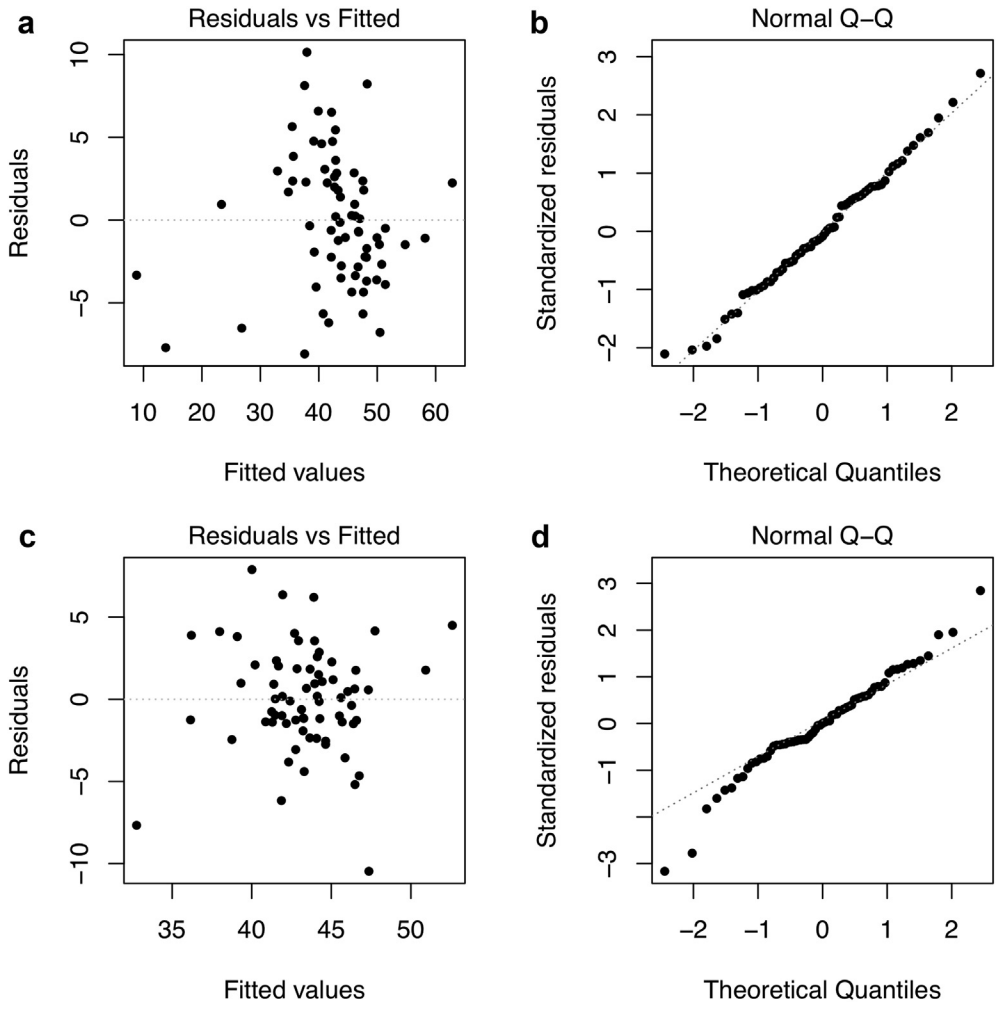

Figure 1 Residual diagnostic plots for (a, b) brainstem, and (c, d) spinal cord doses to 2% brainstem volume. Residual versus fitted plot shows linear relationship in variables and homoscedasticity in residuals and normal Q–Q plot shows normality of residuals.

for patient 42, the minimum distance was >2.8 SD above the mean. For patient 4, the minimum distance was 0 owing to volume overlap between the brainstem and PTV. Because this minimum distance is almost 2 SD below the mean, the close proximity to the PTV contributed to the high brainstem D_2.

In the case of patient 1, the minimum distance was >1.7 SD below the mean. The close proximity to the PTV contributed to a high spinal cord D_2. Conversely, for patient 43, the minimum distance was >3.7 SD above the mean, which contributed to the low spinal cord D_2. Patient 19's spinal cord spinal cord D_2 remained in control in the conventional chart, but the risk-adjusted chart signals OC after accounting for the risk factors. For this patient, the larger PTV (1.7 SD above the mean) and higher PTV dose (above average PTV D_{60} and D_{80}) was used to adjust the spinal cord D_2 to be unusually low in the risk-adjusted chart. Similar observations were made from other best fit risk factor variables that directly led to the adjustments between the conventional and risk-adjusted control charts for the remaining D_x on the brainstem, spinal cord, ipsilateral parotid, and average monitor units.

At clinics, the risk-adjusted control chart can potentially improve the workflow and quality of radiation treatments. For example, the proposed chart can evaluate an optimized plan to be of high quality (in control) after accounting for risk factors, even if the OAR planning criteria are not met. In other words, the control chart can identify plans where improving the OAR planning criteria is only possible at the cost of worsening other risk factors, such as tumor dose distribution. Thus, the proposed chart can help reduce planning time and improve quality for such cases by identifying truly high-quality plans without having to iteratively modify planning criteria.

Furthermore, the typical treatment planning approval process in a clinical setting involves comparing maximum dose and DVH criteria for the OARs to established guidelines, such as Quantitative Analyses of Normal Tissue Effects in the Clinic.[23] These guidelines are

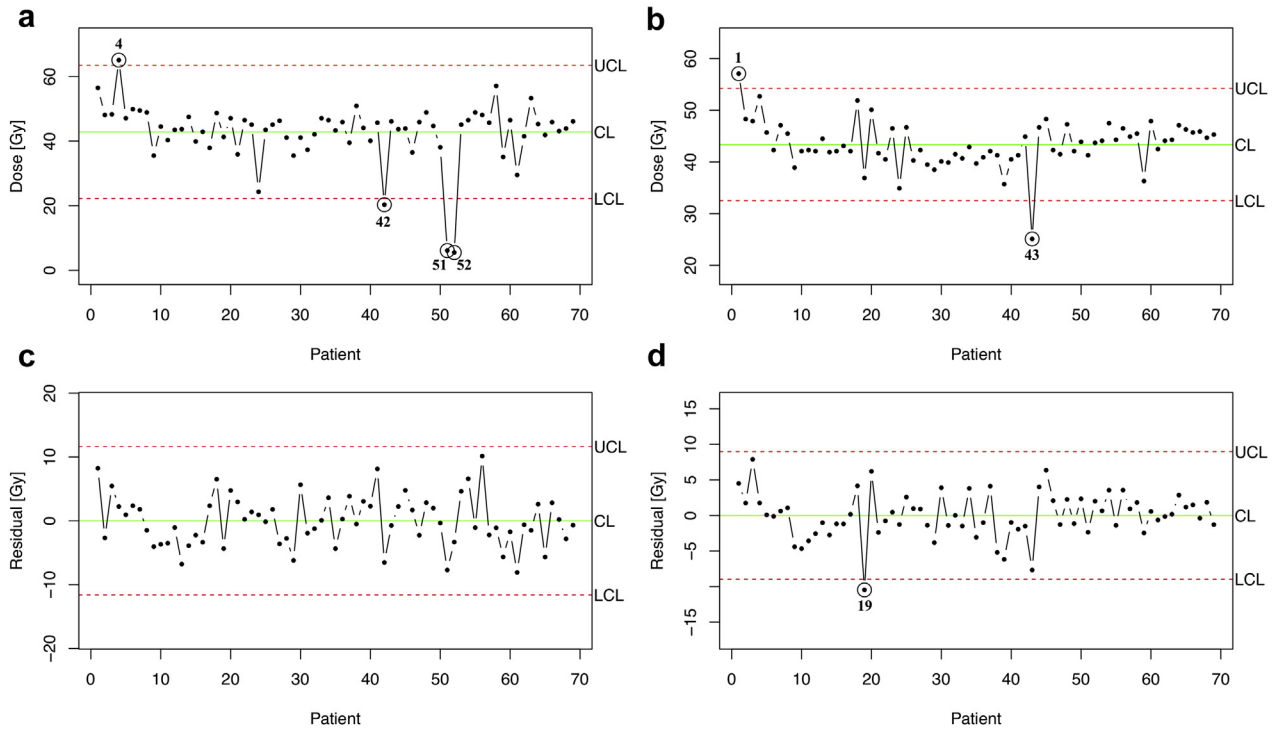

Figure 2 (a) Conventional and (b) risk-adjusted control charts for brainstem doses to 2% brainstem volume and (c) conventional and (d) risk-adjusted control charts for spinal cord doses to 2% brainstem volume.

directly related to risk of toxicity. Of note, a proxy for plan quality, as discussed herein, is not necessarily correlated to toxicity probability. The proposed quality control methodology aims to identify unusual plans based on historical patient data and is independent of meeting toxicity constraints. For patients who need to be treated for secondary cancers or re-treated for a primary cancer, the reduction of low doses that already meet the toxicity guidelines will be of value. An unusually low OAR dose is of less concern, but an OAR dose that exceeds the UCL does increase the potential for adverse effects and can be a cause for an additional review of the plan.

The proposed chart can also be used to monitor future patient's plans rather seamlessly. Given an existing database of already treated patient plans, the risk adjustment regression Equation 1 can be used to model any plan quality proxy, such as an OAR D_x. Next, residuals obtained from the adjustment model can be used to establish control limits using Equations 2 and 3. Similarly, for a new patient, Equation 1 can be used to model for the same quality proxy, and Equations 2 and 3 can be used to compute the residual. Finally, the residual is plotted on the risk-adjusted control chart, reflecting the quality of the new patient's treatment plan. If the residual is within the limits, the plan is acceptable. However, if the residual is OC, further review is needed. Another appealing feature of a control chart is the ability to capture the trend of the planning parameter over time because patients are plotted in chronological order. Not only can trends in planning criteria be visually observed, there are additional order-based quality evaluation criteria that accompany the control limits, which can be employed in the risk-adjusted control chart setting (out of the scope of this work). These include 8 consecutive points above or below the centerline, 4 of 5 consecutive points beyond ± 1 SD, and 2 of 3 consecutive points beyond ± 2 SD that signal a plan to be unusual. These additional measures of quality over time do not exist and cannot be accomplished within the current KPB framework.

In addition to the anatomic parameters discussed herein, other parameterizations, such as portions of OAR outside the primary target fields and tumor site, can be easily considered.[6,11] Beam information, such as beam angles and field arrangements, can also be included. However, given that there are many such planning parameters, external factors, and anatomic parametrizations that could be included in this framework, an exhaustive model considering all risk factors is outside of the scope of this proof of concept.

Unlike control charts that monitor mean, range, or SD that use large subgroup sizes and rely on the central limit theorem, individual charts include a subgroup size of one and thus are sensitive to nonnormal distributions.[18] Nonparametric control charts have been studied in the past, and change-point methods have been proposed to detect shifts in observations over a period in time.[24,25]

Although effective to detect prolonged shifts, these approaches are limited in identifying sporadic OC points and do not account for risk factors that may affect the process quality. As a result, developing nonparametric risk-adjusted control charts serve as an extension of this work. In the same context, extreme outliers can affect the width of the control limits when SD is used to capture the variability in the data. In such a case, a more robust estimate of variability, such as median absolute deviation, can be used to compute the control limits.[26] Furthermore, standardizing the contouring of OARs can reduce outliers, leading to more acceptable control limits. Moreover, the quality control of OAR DVH can also be extended to a multivariate setting or profile monitoring, where multiple OAR D_x or the entire OAR DVH profile can be parametrized for risk adjustment. Capability indices have been used by Breen et al[27] and Sanghangthum et al[28] in radiation therapy to demonstrate that their dosimetry process was within acceptable user specification limits. Extending the indices to the risk-adjusted setting will be considered in future work.

Ideally, clinics should have the ability to readily apply the control chart framework to their patient data without tediously curating and formatting the data, as was required for this proof of concept. For example, DVH Analytics is an open-source DICOM radiation therapy database software that can read large batches of DICOM files and compute various anatomic and dose values, store them in an SQL database, and provide a graphical user interface to explore the data.[29]

Furthermore, as the treatment planning process continues to evolve toward fully automated planning, the quality assurance of planning criteria becomes even more important because optimization algorithms often produce locally optimum solutions in radiation therapy.[30] The proposed control chart methodology can employ large databases of existing plans to help identify locally optimal plans that could be improved through reoptimization of planning parameters.

Conclusions

A risk-adjusted control chart was developed to detect unusual planning criteria after accounting for patient- and treatment-specific risk factors, allowing for multiple plan comparisons on a large scale. When evaluating 69 head and neck cases, the risk-adjusted control charts were robust to the interpatient variations in anatomy, tumor dose prescription, use of chemotherapy, and surgery. Furthermore, the charts reduced the number of unusual plans compared with conventional control charts. Detecting unusually high OAR doses in risk-adjusted charts provides the clinician with a chance to review the treatment plan within the context of historical patient data and probe for possible improvements to reduce the potential for adverse effects.

Supplementary data

Supplementary material for this article can be found at https://doi.org/10.1016/j.adro.2019.11.006.

References

1. Das IJ, Andersen A, Chen ZJ, et al. State of dose prescription and compliance to international standard (ICRU-83) in intensity modulated radiation therapy among academic institutions. *Pract Radiat Oncol*. 2017;7:e145-e155.
2. Chung JB, Kim JS, Ha SW, Ye SJ. Statistical analysis of IMRT dosimetry quality assurance measurements for local delivery guideline. *Radiat Oncol*. 2011;6:27.
3. Pawlicki T, Whitaker M. Variation and control of process behavior. *Int J Radiat Oncol Biol Phys*. 2008a;71:S210-S214.
4. Pawlicki T, Yoo S, McMillan SK, et al. Process control analysis of IMRT QA: Implications for clinical trials. *Phys Med Biol*. 2008b;53:5193.
5. Gerard K, Grandhaye JP, Marchesi V, Kafrouni H, Husson F, Aletti P. A comprehensive analysis of the IMRT dose delivery process using statistical process control (SPC). *Med Phys*. 2009;36:1275-1285.
6. Palaniswaamy G, Brame RS, Yaddanapudi S, Rangaraj D, Mutic S. A statistical approach to IMRT patient-specific QA. *Med Phys*. 2012;39:7560-7570.
7. Nordström F, af Wetterstedt S, Johnsson S, Ceberg C, Bäck SÅ. Control chart analysis of data from a multicenter monitor unit verification study. *Radiother Oncol*. 2012;102:364-370.
8. Moore KL, Brame RS, Low DA, Mutic S. Experience-based quality control of clinical intensity modulated radiotherapy planning. *Int J Radiat Oncol Biol Phys*. 2011;81:545-551.
9. Zhu X, Ge Y, Li T, Thongphiew D, Yin FF, Wu QJ. A planning quality evaluation tool for prostate adaptive IMRT based on machine learning. *Med Phys*. 2011;38:719-726.
10. Appenzoller LM, Michalski JM, Thorstad WL, Mutic S, Moore KL. Predicting dose-volume histograms for organs-at-risk in IMRT planning. *Med Phys*. 2012;39:7446-7461.
11. Yuan L, Ge Y, Lee WR, Yin FF, Kirkpatrick JP, Wu QJ. Quantitative analysis of the factors which affect the interpatient organ-at-risk dose sparing variation in IMRT plans. *Med Phys*. 2012;39:6868-6878.
12. Lian J, Yuan L, Ge Y, et al. Modeling the dosimetry of organ-at-risk in head and neck IMRT planning: An intertechnique and interinstitutional study. *Med Phys*. 2013;40:121704.
13. Wu B, Ricchetti F, Sanguineti G, et al. Patient geometry-driven information retrieval for IMRT treatment plan quality control. *Med Phys*. 2009;36:5497-5505.
14. Ge Y, Wu QJ. Knowledge-based planning for intensity-modulated radiation therapy: A review of data-driven approaches. *Med Phys*. 2019;46:2760-2775.
15. Berry SL, Ma R, Boczkowski A, Jackson A, Zhang P, Hunt M. Evaluating intercampus plan consistency using a knowledge based planning model. *Radiother Oncol*. 2016;120:349-355.
16. Shewhart WA. The application of statistics as an aid in maintaining quality of a manufactured product. *J Am Stat Assoc*. 1925;20:546-548.
17. Grigg O, Farewell V. An overview of risk-adjusted charts. *J R Stat Soc Ser A Stat Soc*. 2004;167:523-539.
18. Montgomery DC. *Statistical quality control*. 7th ed. New York, NY: Wiley; 2012.

19. Deasy JO, Blanco AI, Clark VH. CERR: A computational environment for radiotherapy research. *Med Phys*. 2003;30:979-985.
20. Roy A, Das IJ, Nohadani O. On correlations in IMRT planning aims. *J Appl Clin Med Phys*. 2016;17:44-59.
21. Fiorino C, Dell'Oca I, Pierelli A, et al. Significant improvement in normal tissue sparing and target coverage for head and neck cancer by means of helical tomotherapy. *Radiother Oncol*. 2006;78:276-282.
22. Pregibon D, Hastie TJ. Generalized linear models. In: *Statistical Models in S*. Abingdon, United Kingdom: Routledge; 2017. pp.195-247.
23. Marks LB, Yorke ED, Jackson A, et al. Use of normal tissue complication probability models in the clinic. *Int J Radiat Oncol Biol Phys*. 2010;76:S10-S19.
24. Ning W, Yeh AB, Wu X, Wang B. A nonparametric phase I control chart for individual observations based on empirical likelihood ratio. *Qual Reliab Eng Int*. 2015;31:37-55.
25. Sullivan JH, Woodall WH. A control chart for preliminary analysis of individual observations. *J Qual Technol*. 1996;28:265-278.
26. Adekeye KS, Azubuike PI. Derivation of the limits for control chart using the median absolute deviation for monitoring nonnormal process. *J Math Statist*. 2012;8:37-41.
27. Breen SL, Moseley DJ, Zhang B, Sharpe MB. Statistical process control for IMRT dosimetric verification. *Med Phys*. 2008;35:4417-4425.
28. Sanghangthum T, Suriyapee S, Srisatit S, Pawlicki T. Statistical process control analysis for patient-specific IMRT and VMAT QA. *J Radiat Res*. 2012;54:546-552.
29. Cutright D, Gopalakrishnan M, Roy A, Panchal A, Mittal BB. DVH Analytics: A DVH database for clinicians and researchers. *J Appl Clin Med Phys*. 2018;19:413-427.
30. Deasy JO. Multiple local minima in radiotherapy optimization problems with dose–volume constraints. *Med Phys*. 1997;24:1157-1161.