ARTICLE

Predicting radiation treatment planning evaluation parameter using artificial intelligence and machine learning

Frederick Ng¹, Runqing Jiang²³ and James C.L. Chow¹³*

¹ Department of Physics, Ryerson University, Toronto, ON M5B 2K3, Canada
² Department of Medical Physics, Grand River Regional Cancer Centre, Grand River Hospital, Kitchener, ON N2G 1G3, Canada
³ Department of Physics, University of Waterloo, Waterloo, ON N2L 3G1, Canada
⁴ Radiation Medicine Program, Princess Margaret Cancer Centre, University Health Network, Toronto, ON M5G 1X6, Canada
⁵ Department of Radiation Oncology, University of Toronto, Toronto, ON M5T 1P5, Canada

E-mail: james.chow@rmp.uhn.ca

Keywords: treatment planning evaluation, artificial intelligence, machine learning, data analysis, prostate VMAT, dose-volume histogram

Abstract

Purpose: This study suggested a new method predicting the dose-volume parameter for radiation treatment planning evaluation using machine learning, and to evaluate the performance of different learning algorithms in the parameter prediction. Methods: Dose distribution index (DDI) for fifty prostate volumetric modulated arc therapy plans were calculated, and compared to results predicted by machine learning using algorithms, namely, linear regression, tree regression, support vector machine (SVM) and Gaussian process regression (GPR). Root mean square error (RMSE), prediction speed and training time were determined to evaluate the performance of each algorithm. Results: From the results, it is found that the square exponential GPR algorithm had the smallest RMSE, relatively high prediction speed and short training time of 0.0038, 4,100 observation/s and 0.18 s, respectively. All linear regression, SVM and GPR algorithms performed well according to their RMSE in the range of 0.0038–0.0193. However, RMSE of the medium and coarse tree regression algorithms were found larger than 0.03, showing that they are not suitable for predicting DDI in this study. Conclusion: Machine learning can be used to predict dose-volume parameter such as DDI in radiation treatment planning QA. Selection of a suitable machine learning algorithm is important to determine the parameter effectively.

Introduction

In radiotherapy, a dose delivery plan is created by radiation dosimetrist and medical physicist. This dose plan contains information about how to irradiate a volume-of-interest using a dose delivery technique. To guarantee a high conformal dose deposited at the target, while simultaneously sparing the nearby organs-at-risk (OARs), a quality assurance (QA) program is needed to evaluate the dose plan to ensure it fulfills the minimum standard [1–3]. For the convenience of plan evaluation, dose distribution in the irradiation volume such as target or OAR is expressed as dose-volume histograms (DVHs). DVH is a histogram relating absorbed dose to the irradiated volume in treatment planning, and is useful for the calculation of different plan evaluation parameters [4–6]. These dose-volume and radiobiological parameters are based on medical physics, radiobiology and mathematical expression reflecting the dose distribution merit of a treatment plan in QA [7–11].

Dose distribution index (DDI) is a relatively new dose-volume evaluation parameter proposed to help decision-making in radiation treatment planning QA [12]. This index consolidates the dose coverage conformity and the homogeneity for the planning target volume (PTV), OAR, and remaining target-at-risk (RVR) into one value. DDI is proved to be an effective and straightforward way to evaluate treatment plan in radiotherapy delivery program [12, 13]. However, calculation of DDI is still complicated as it involves a number of DVHs from the targets and OARs. This is an issue especially when there are many potential plans generated for the treatment which require a fast evaluation [14, 15]. Since there are already many studies on dose
prediction for treatment planning using machine learning algorithms [16–18], we solved this problem to predict the DDI value using machine learning with training on related DVHs to predict the DDI value [19–21].

This Science Note focused on machine learning applications after the dose and DVH calculation in radiation treatment planning. The aim is to confirm that machine learning algorithms can efficiently predict dose-volume parameters such as DDI in treatment planning QA. Prostate volumetric modulated arc therapy (VMAT) plans were used in the machine learning [21, 22].

Methods

Dose distribution index

The DDI is written as [12]:

\[
DDI = \frac{1}{3} (W_T I_T + W_O I_O + W_R I_R),
\]

where \(W_T\), \(W_O\), and \(W_R\) are the weights assigned to the PTV, OAR, and RVR and \(I_T\), \(I_O\), and \(I_R\) are the components that quantify the plan quality on the PTV, OAR, and RVR.

The components defined the dose coverage and conformity of the PTV \((I_T)\), OAR \((I_O)\) and RVR \((I_R)\) in equation (1) are defined as:

\[
I_T = 1 - \frac{1}{1} \left( 1 - \left( \frac{\int_0^{D_{MT}} V_T(D) dD}{D_{PTV}} \right) \left( \frac{D_{MT}}{D_{PTV}} \right) \right),
\]

\[
I_O = \frac{1}{N} \sum_{i=1}^{N} W_{O_i} \left( 1 - \left( \frac{\int_0^{D_{MT}} V_O(D) dD}{D_{OAR}} \right) \left( \frac{D_{MT}}{D_{OAR}} \right) \right),
\]

and

\[
I_R = \left( 1 - \left( \frac{\int_0^{D_{MT}} V_R(D) dD}{D_{RVR}} \right) \left( \frac{D_{MT}}{D_{RVR}} \right) \right),
\]

where \(\int V_T(D) dD, \int V_O(D) dD\) and \(\int V_R(D) dD\) represent the area under the PTV, OAR and RVR in the DVH curves, respectively. In equation (2), PTV is the planning target volume, \(D_p\) is the prescribed dose, \(D_m\) is the maximum dose received completely and \(D_{MT}\) is the maximum dose received within the PTV. In equations (3) and (4), N is the number of OAR, and \(W_{O_i}\) is a non-negative weight that reflects the relative clinical importance of the \(i^{th}\) OAR.

DDI can be determined solely from the prescribed dose and DVHs of the PTV, OAR, and RVR involved [12]. All this information is usually available from the computerized treatment planning system. The influence of the PTV, OAR, and RVR in conformity and homogeneity of these volumes are estimated and then combined in a weighted sum. The weights are set to correspond to the volumes’ relative importance assigned in the radiation treatment planning process. Instead of evaluating a treatment plan based on calculating the corresponding evaluation parameters from DVHs of the targets and OAR [7, 9, 10], DDI simplified the QA process by providing only one value based on a group of DVHs. However it can be seen that to determine the DDI, all the corresponding DVHs of the targets, OARs and RVRs in the treatment plan have to be considered. Since there are typically over 500–800 dose-volume points in one DVH and there are generally 5–10 DVHs to be considered in a DDI calculation, it is worthwhile to investigate if it is more efficient to use machine learning to predict the DDI.

In this study, fifty prostate VMAT plans from the Grand River Hospital were used. The dual-arc technique was used with the 6 MV photon beams [23]. DVHs of the targets and OARs in each plan were determined to calculate the DDI as per equation (1).

Workflow of machine learning

MATLAB’s regression learner in the Statistical and Machine Learning toolbox App was used in this study. The workflow of machine learning is shown in figure 1.
Data collection and preprocessing

The radiation treatment plan data consisted of the DVH of the PTV, and each of the OAR. The OAR of these plans included the rectum, bladder, left femur and right femur. In data preprocessing, features of DVH were transformed to the maximum dose received by a certain percent of relative volume, defined as $D_x$. The $x$ represents the percentage of the relative volume that receives a maximum dose of $D_x$. The total amount of features required was tested iteratively and it is found that $D_{100\%}$, $D_{75\%}$, $D_{50\%}$, $D_{25\%}$ and $D_M$ as shown in figure 2 were adequate and efficient to predict the DDI value. $D_M$ is the maximum dose of the DVH.

Model selection, training and evaluation

To select the appropriate models to represent the response (figure 1), it is found that supervised regression machine learning algorithms are suitable [24]. The algorithms included linear regression [25], tree regression [26], SVM [27, 28] and GPR [29] used in this study. They were examined for their performance in predicting the DDI value. In the training process, $k$-folds cross validation was carried out [30]. In $k$-fold cross validation the dataset is divided into $k$ groups. Each group takes a turn being the hold out fold while the other $k - 1$ groups acts as the training data. The total performance is the average of all folds. $k$-fold cross validation is useful, as a subset of data does not need to be held out, making it particularly beneficial for a small training dataset. A 4-fold cross validation was used in this study. The root mean square error (RMSE) was used to evaluate the performance of all
machine learning algorithms. In addition, the prediction speed (i.e. number of observation per second) and training time for each algorithm were recorded for comparison.

Results

An actual versus predicted plot of the DDI values using different machine learning models trained with the optimal number of features was used to identify trends of the prediction. Results of actual versus predicted plots
for the linear regression, tree regression (fine, medium and coarse), SVM (linear, quadratic and cubic) and GPR (exponential, matern 5/2, rational quadratic, and square exponential) are shown in figures 3–6, respectively.

The RMSE, R-squared, prediction speed and training time for the linear regression, tree regression, SVM and GPR models are shown in table 1. They are ordered from the best performance to the worst.

Discussion

From the results in figures 3–6 and table 1, it is seen all machine learning algorithms examined reliably predict the DDI except the tree regression model. The tree regression models (figure 4) were not precise enough for a reliable prediction of DDI. The GPR models (figure 6) were the most accurate among others. It was almost matched the calculated results with RMSE smaller than 0.004 (table 1).

In the linear regression model, the largest coefficient was the $D_{50\%}$ for the PTV. This meant that the $D_{50\%}$ of the PTV was a strong predictor of the DDI. The left and right femur had a large coefficient for the $D_{100\%}$ but the contribution to the overall DDI was small since the $D_{100\%}$ for OAR were small. The tree regression was not accurate enough since the resolution of the tree was not fine enough to distinguish DDI (figure 4). On the other hand, all SVM models performed well (figure 5) with the linear SVM (figure 5(a)) being the best. The SVM models were the most accurate when there were around 5 dose-volume points extracted from each DVH. For the quadratic (figure 5(b)) and linear (figure 5(a)) SVM, the accuracy showed a trend of decreasing accuracy when the number of features was increased. A possible explanation for this was the increasing number of features made it difficult for the SVM to focus on the features that were most predictive. The GPR algorithms created were the most accurate models (figures 6(a)–(d)). The RMSE for the square exponential, matern 5/2, and rational quadratic GPR were very similar as shown in table 1. For the computing time, while our MATLAB program would take about 5 s to determine the DDI from DVHs per treatment plan, it only took less than one second to determine the DDI using the machine learning method. This time advantage is significant when we need to compare hundreds or treatment planning options for a patient.
Conclusions

Machine learning was used to predict the DDI from DVHs of prostate VMAT treatment plans using different models, namely, linear regression, tree regression, SVM, and GPR. It is found that all examined models were precise in its predictions of DDI, except the tree regression model. So it is possible to predict dose-volume parameter from the DVH data using machine learning.
Data availability statement

The data that support the findings of this study are available upon request from the authors.

ORCID iDs

James C L Chow 🐘 https://orcid.org/0000-0003-4202-4855

References

[1] Shafig, J. Barton, M. Noble, D. Lemer C and Donaldson L J 2009 An international review of patient safety measures in radiotherapy practice Radiother and Oncol 92 15–21
[2] Ishikura S 2008 Quality assurance of radiotherapy in cancer treatment: toward improvement of patient safety and quality of care Jap J Clin Oncol 38 723–9
[3] Chow J C L, Jiang R and Kiçiaa A 2016 Dose-volume consistency and radiobiological characterization between prostate IMRT and VMAT plans Int J Cancer Ther Oncol. 4 447
[4] Grigorov G N, Chow J C L, Grigorov L, Jiang R and Barnett R B 2006 IMRT: improvement in treatment planning efficiency using NTCP calculation independent of the dose-volume-histogram Med. Phys. 33 1250–8
[5] Chow J C L, Markel D and Jiang R 2010 Calculation of normal tissue complication probability using Gaussian error function model Med. Phys. 37 4924–9
[6] Chow J C L, Runqing J and Daniel M 2009 The effect of interfraction prostate motion on IMRT plans: a dose-volume histogram analysis using a Gaussian error function model Appl Clin Med Phys. 10 79–95
[7] Chow J C L, Jiang R, Kiçiaa A and Daniel M 2016 Dosimetric comparison between the prostate intensity-modulated radiotherapy (IMRT) and volumetric-modulated arc therapy (VMAT) plans using the planning target volume (PTV) dose-volume factor Radiother in Pract. 15 263–8
[8] Deschavanne P J and Fertil B 1996 A review of human cell radiosensitivity in vitro Int J Radiat Oncol Biol Phys 34 251–66
[9] Feuvret L, Noël G, Mazeron J and Bey P 2006 Conformity index: a review Int J Radiat Oncol Biol Phys. 64 333–42
[10] Kataria T, Sharma K, Subramani V, Karrthick K P and Bish S S 2012 Homogeneity index: an objective tool for assessment of conformal radiation treatments J Med Phys. 37 207–13
[11] Baumann M and Petersen C T 2005 TCP and NTCP: a basic introduction Rays. 30 99–104 PMID:16294901
[12] Alonso J C, Herrera M A and Nunez L 2015 A dose-volume histogram based decision-support system for dosimetric comparison of radiotherapy treatment plans Radiat Oncol. 10 1263
[13] Chow J C L, Jiang R and Lu X 2019 Evaluation of plan optimizers in prostate VMAT using the dose distribution index Radiother in Pract. 18 323–8
[14] Hussein M, Heijmen B J, Verellen D and Nisbet A 2018 Automation in intensity modulated radiotherapy treatment planning—a review of recent innovations BJR. 91 20180270
[15] Kravenbuhl J, Zambrurni M, Ghandour S, Pachoud M, Lang-Tanadini S, Tol J, Guckenberger M and Verbakel W F 2018 Planning compliance of five automated treatment planning solutions for locally advanced head and neck cancer Radiat Oncol. 13 170
[16] Valdes G, Simone C B II, Chen J, Lin A, Yom S S, Pattison A J, Carpenter C M and Solberg T D 2017 Clinical decision support of radiotherapy treatment planning: a data-driven machine learning strategy for patient-specific dosimetric decision making J Radiat Oncol. 125 392–7
[17] Jones S, Hargrave C, Deegan T, Holt T and Mengersen K 2020 Comparison of statistical machine learning models for rectal protocol compliance in prostate external beam radiation therapy Med. Phys. (https://doi.org/10.1002/mp.14044)
[18] Kontaxi C, Bol G H, Legrand J J and Raamakers B W 2020 Deep dose: towards a fast dose calculation engine for radiation therapy using deep learning Phys. Med. Biol. (https://doi.org/10.1088/1674-1137/ab7630)
[19] Chow J C L 2017 Internet-based computer technology on radiotherapy Rep Pract Oncol Radiat. 22 455–62
[20] Domingos P 2012 A few useful things to know about machine learning Commun of the ACM. 55 78
[21] Schwarzer G, Vach W and Schumacher M 2000 On the misuses of artificial neural networks for prognostic and diagnostic classification in oncology Stat Med 19 541–61
[22] Isa M, Jiang R, Kiçiaa A, Rehman J, Afzal M and Chow J C L 2016 Dosimetric and radiobiological characteristics of prostate IMRT and VMAT: a single-institution review of 90 cases J Med Phys. 41 162–8
[23] Chow J C L and Jiang R 2013 Prostate volumetric-modulated arc therapy: dosimetry and radiobiological model variation between the single-arc and double-arc technique J Appl Clin Med Phys. 14 3–12
[24] Kotsiantis S B, Zaharakis I and Pintelas P 2007 Supervised machine learning: a review of classification techniques Emerging Artificial Intelligence.Applications in Computer Engineering. 160 3–24
[25] Matloff N 2017 Statistical Regression and Classification (Boca Raton, Florida, United States: CRC Press) pp 65–70
[26] Loh W 2011 Classification and Regression Trees (Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery) 1, pp 14–23
[27] Suykens J A and Vandewalle J 1999 Least squares support vector machine classifiers Neural Proc lett. 9 293–300
[28] Tomioka R, Suzuki T and Sugiyama M 2011 Super-linear convergence of dual augmented Lagrangian algorithm for sparsity regularized estimation Mach Learning Res. 12 1357–86
[29] Seeger M 2004 Gaussian processes for machine learning Int J Neural Sys. 14 69–106
[30] Bengio Y and Grandvalet Y 2004 No unbiased estimator of the variance of k-fold cross-validation Mach Learning Res. 5 1089–105