Aus der Klinik für Strahlentherapie und Radioonkologie
der Medizinischen Fakultät Mannheim, Universität Heidelberg
(Kommissarischer Direktor: Dr. med. Michael Ehmann)

A KNOWLEDGE-BASED QUANTITATIVE APPROACH TO CHARACTERIZE TREATMENT PLAN QUALITY: APPLICATION TO PROSTATE VMAT PLANNING

Inauguraldissertation
zur Erlangung des Doctor scientiarum humanarum (Dr. sc. hum.)
der
Medizinischen Fakultät Mannheim
der Ruprecht-Karls-Universität
zu
Heidelberg

vorgelegt von
Buthayna Alnaalwa

aus
Tulkarm, Palästina
2022

Dekan: Herr Prof. Dr. med. Sergij Goerdt
Referent: Herr Prof. Dr. med. Frederik Wenz

TABLE OF CONTENTS

Page

LIST OF FIGURES ..... **I**

LIST OF TABLES ..... **II**

LIST OF ABBREVIATIONS ..... **III**

1 INTRODUCTION **1**

1.1 Overview of the treatment planning process *1*

1.2 Knowledge-based radiation therapy *2*

1.3 Motivation *4*

1.4 Aims *4*

2 MATERIALS AND METHODS **5**

2.1 Ethics approval and Data base of plans *5*

2.2 Extraction of information from the treatment plans *5*

2.3 Data analysis *6*

2.3.1 The mean dose-at-distance function 6

2.3.2 The reference set versus query set 6

2.3.3 Plan evaluation metrics: $M_{q,r}$ value 7

2.3.4 Statistical Analysis *8*

2.3.5 Model Validation *8*

2.3.6 Refinement of database *9*

3 RESULTS **11**

4 DISCUSSION **20**

5 SUMMARY **25**

6 REFERENCES **25**

7 CURRICULUM VITAE **31**

8 ACKNOWLEDGEMENT **33**

LIST OF FIGURES

Figure 1. A schematic representation of the study workflow. Homogenous and SIB VMAT TPs form the data base of this study 10

Figure 2(a) Probability ellipses of a bivariate normal distributed Mq,bladder versus Mq,rectum for homogenous VMAT prostate plans. 12

Figure 3 (a) Probability ellipses of a bivariate normal distributed $M_{q,bladder}$ versus $M_{q,rectum}$ for SIB VMAT prostate plans. 13

Figure 4. Validation of the model of the bladder and rectum of the homogenous plans. 17

Figure 5. Validation of the model of the bladder and rectum of the SIB plans. 18

LIST OF TABLES

Table 1. $M_{q,r}$ values of bladder and rectum, before and after re-optimization for homogenous VMAT prostate plans. 15

Table 2. $M_{q,r}$ values of bladder and rectum, before and after re-optimization for SIB VMAT prostate plans. 16

Table 3. Summary of $M_{q,r}$ values - Mean ± Standard Error (SE), STD ± SE, the correlation parameter between bladder and rectum $M_{q,r}$ values (rho factor) - for both treatment groups before and after the data-base was refined. 19

LIST OF ABBREVIATIONS

PTV	Planning Target Volume
OAR	Organ At Risk
DVC	Dose-Volume Constraint
NTCP	Normal Tissue Complication Probability
TP	Treatment Plan
VMAT	Volumetric Modulated Arc Therapy
KBRT	Knowledge-Based Radiation Therapy
OVH	Overlapped Volume Histogram
DVH	Dose Volume Histogram
DTH	Distance to Target Histogram
MCO	Multi-Criteria Optimization
QC	Quality Control
SIB	Simultaneous Integrated Boost
CERR	Computational Environment for Radiotherapy Research

1 INTRODUCTION

1.1 Overview of the treatment planning process

The overall aim of the treatment planning process is to translate the therapeutic requirement of the patient into a set of treatment instructions that will enable treatment while minimizing the risk of complications from radiation. Treatment planning typically starts with (1) images of the treatment volume (e.g., from CT or MRI scans) and, (2) the desired absorbed dose of radiation which is to be delivered to the planning target volume (PTV), and (3) the maximum absorbed dose which can be safely absorbed by tissue structures, such as organs at risks (OARs) which are located within the treatment volume adjacent to the target volume. Both the PTV and any OAR can have complex three-dimensional shapes, complicating the creation of a treatment plan.[1,2] The optimum treatment plan is thus that which achieves the prescribed target dose at a minimum of the organ at risks doses.[3,4]

The success of radiation therapy depends on the ability to deliver the proper amount of radiation to cancerous cells while protecting healthy tissues. Thus, in treatment planning, the desired target and healthy tissue dose distribution are often specified as dose-volume constraints (DVCs), which define both how much radiation is required in the target volume, as well as the limits on radiation to the OARs.[5,6] More specifically, the DVCs may specify that a tumour should receive an absorbed dose of at least X Gy in y% of the tumour volume. The DVCs are used to define the plan objectives, but the definition of the objective function is arguably the most challenging part of plan optimization. The objective function may reflect trade-offs arising while trying to achieve a very good compromise between the desired absorbed dose in the target while keeping the normal tissue complication probability (NTCP) as low as possible.[7,8] The proper trade-off between these competing goals needs the planner to set and adapt the dose objectives for the targets and OARs to fit dose volume constraints using an iterative approach[9] in the inverse treatment planning process. Running more iterations to improve the treatment plan (TP) can be counterproductive if unachievable constraints are used. The most common dose volume constraints have typically been determined using the Radiation Therapy Oncology Group protocol (RTOG) recommendations.[10]

The process of generating a clinical treatment plan is generally an iterative time-consuming process. Since it involves optimizing large numbers of variables and complex matrix manipulations to calculate new treatment plans, it requires a relatively long substantial computational time even when using modern high-speed computers. These problems are even more difficult in treatment planning for arc therapy like volumetric intensity-modulated arc therapy (VMAT), which uses a moving source of radiation (the gantry of the accelerator may rotate along one or more arcs and deliver radiation continuously).[11-13] The greater degree of freedom which is provided by such complex radiotherapy techniques has made the task of developing treatment plans even more difficult.[7,13] Although these systems provide high-quality plans, the optimal solution generally depends on the planner's time and experience to convert the clinical goals into an optimized plan accounting for the relation between the PTV coverage and the OARs sparing.

1.2 Knowledge-based radiation therapy

Knowledge-based radiation therapy (KBRT) is an approach to improve efficiency and to reduce variability in treatment planning.[14-17] A variety of KBRT approaches have been reported. Some KBRT approaches are based on search and retrieval of prior optimized patient plans to guide the generation of new plans based on the similarities of their geometries.[14,15,18-20] For example, Wu et al.[15] retrieved a prior patient plan based on similarities of overlapped volume histograms (OVHs), to guide the optimization of the new plan. Chanyavanich et al.[18] used the beam's eye view of the treatment region of a new case to find the patient with the most similar anatomy from a database of previously treated patients. The planning constraints of the matching plan were used to define the planning goals of the new plan. Another approach predicts the dose volume histograms (DVHs) of the OARs based on their distance to target histograms (DTHs).[21,22] More specifically, in this approach, a quantitative evaluation tool is proposed based on machine learning to characterize the relation between the DVH and anatomical features for each patient plan. This information could be used to estimate the DVH for a new TP or to optimize plan quality.[23,24]

Various efforts using these KB approaches were presented to improve treatment planning consistency,[24-27] including RapidPlan[25] in Eclipse, Multi-Criteria Optimization (MCO) in RaySearch.[26] RapidPlan is a commercially available DVH-guidance approach. In this approach a mathematical model that correlates the geometrical features of patients to previously achieved dosimetry is used to derive the achievable DVHs for prospective cases.[28] The key limitation of DVH-based approaches is the lack of spatial information. Planners may need extra work to deal with a case with uncommon OAR/target geometry. Furthermore, the DVH only predicts the delineated regions. Optimizing the dose to improve the conformality of tissues outside the delineated region might not be considered. Another approach called MCO tries to find the optimal trade-offs between target coverage and normal tissues sparing. Thus, it helps less-experienced planners to produce high-quality IMRT plans. In MCO, a pareto surface containing a spectrum of fitting plans is automatically generated by the optimizer. Then, planners navigate through the combinations based on specified trade-offs objectives to choose a pareto-optimal plan.[29,30] However, pareto-optimal plans are not clinically optimal plans and can be clinically highly undesirable. Pareto optimal is still the best clinically acceptable plan in MCO. The limitation of this approach is the large number of automatically generated pareto-optimal plans investigating many parameters to optimize the different clinical objectives. Generating a large number of optimal plans needs intensive computation resources, in addition to the accompanied difficulties in the selection of an optimal plan. However, using prior knowledge to automatically predict the DVH of the new plan cannot provide any reasonable estimation of the achievable absorbed dose distributions among different patients.[31]

Another KBRT approach described by Nwankwo et al.[32] predicts the dose at voxel level of the OARs. More specifically, this algorithm relates the OARs dose of the voxels to their geometric proximity to the PTV. This algorithm was derived from the analysis of a database of high-quality TPs. As a result, this method objectively considers patient-specific anatomical variations in the treatment planning process, and thus enables the comparison of plans irrespective of patient anatomical geometries. Furthermore, the algorithm can predict the 3D absorbed dose distribution in the organ of interest.[32]

1.3 Motivation

Even though the inverse treatment planning process is highly computerised, the process still contains an iterative approach to convert clinical requirements into DVH based objectives. A High level of planner intervention is required to produce a high-quality treatment plan, as the planner progressively improves the plan until it becomes clinically acceptable. In each optimization round, the DVH objectives should be optimized based on the planner subjective decision and clinical knowledge. Thus, planner can expend days to generate/improve new treatment plans.[9] Typically, optimizing goals are determined from population-based data, or the RTOG guidlines and clinical feedbacks. These population-based recommendations can be meaningful for some plans, but not for others, due to patients' geometric variations. However, based on the DVH objectives, the generated plan quality might be, not only, highly variable between planners, but also can be far from being the best optimal plan, if the dose to the normal tissue is not minimized to the best extent possible.[33] Thus, plans at radiation therapy centres with limited experience may not be clinically optimal.[16,34] To address this need, we proposed a quality control (QC) method to estimate the achievable OARs dose sparing. Such a QC algorithm should not entirely rely on personal judgment but should consider the geometric variations among patients. Thus, this study is motivated to provide guidance of the best achievable OAR sparing for each specific patient anatomy, and to eliminate the trial and error approach of defining the objective function during the treatment planning process.

1.4 Aims

To characterize treatment plan quality, a quantitative QC tool based on the method published by Nwankwo et al.[32] is developed and proposed. The tool is validated using VMAT prostate plans by estimating a threshold to spot the sub-optimal plans to achieve further dose sparing in rectum and bladder using knowledge learned from prior plans. More specifically, the dose of the binned voxels according to their distant location from the PTV surface between the query plan and the reference plans were compared. Thus, the difference in the achieved sparing between plans is compared to determine the gain, and the influence of plan objectives on the achieved OAR sparing.

2 MATERIALS AND METHODS

2.1 Ethics approval and Data base of plans

This study was approved by the Medical Ethics Commission of the Medical Faculty Mannheim, Heidelberg University (2015-816R-MA).

Four hundred and fifty (450) VMAT plans that were approved by both physician and physicist for a real prostate cancer treatment at the Department of Radiation Oncology, University Medical Centre, Mannheim, Germany, were used to conduct this study. These VMAT prostate plans were randomly selected from the database of patient plans that were treated between 2007 and 2017. The plans were anonymized for use in this study. The plans were divided into two groups: 181 homogenous prostate VMAT plans, and 269 simultaneous integrated boost (SIB) prostate VMAT plans. The target volumes in the homogenous plans were the prostate and seminal vesicles whereas the target volumes in the SIB plans were the prostate and seminal vesicles and a boost volume. The PTV prescribed doses were variable for the plans. These plans were calculated with Monaco® (CMS, Elekta-Group, UK) that is a fully inverse treatment planning system. The template used for prostate planning[33] specifies the desired PTV coverage and the acceptable doses to the normal tissues. These constraints are adjusted based on trial-and-error planning strategy until the dose distribution is considered to be optimal for the given patient anatomy.

2.2 Extraction of information from the treatment plans

The plans were imported into a MATLAB programming platform using the Computational Environment for Radiotherapy Research (CERR) software.[34] In each patient plan the following variables were extracted for each voxel in the OARs (rectum, bladder):[32,35,36]

1. The coordinates (x, y, z) of each voxel of an OAR. Both OARs (rectum and bladder) were considered in this work.

2. The *dose of each voxel* which was normalized to the prescribed planning target volume (PTV) dose. Dose normalization renders the analysis to be independent of the absolute value of the PTV prescribed dose.

3. The *distance-to-PTV* of each voxel. It is the shortest Euclidean distance between a voxel of an OAR and the surface of the PTV. The negative distance to PTV value is set for the shared voxel between both OAR and PTV. Only the voxel within the beam path was considered in this calculation.
4. Two matrices, one for each OAR (rectum and bladder) were computed from each plan. A row of the matrix represents a voxel while the column specifies its parameters (*dose*, *slice level*, *distance-to-PTV* and the *Cartesian* coordinates).

2.3 Data analysis

2.3.1 The mean dose-at-distance function

The mean *dose-at-distance function* of an organ was calculated from the in-field voxels of the OARs.[32,35,36] The voxels were first divided into bins (0.5 mm bin size according to their distance-to-PTV). The mean dose-at-distance value was calculated for each distance-to-PTV bin, more precisely the mean dose-at-distance values of the in-field voxels were calculated against the centre of the corresponding distance-to-PTV bin.

This function relates the mean dose of all the in-field voxels within a distance-to-PTV bin to the median distance of the voxels to PTV surface. The calculation of this variable is explained in reference [10].

2.3.2 The reference set versus query set

For both groups, homogenous VMAT plans and SIB VMAT plans, the reference sets and the query sets were generated from the same plan matrices of the group. The same (181) homogenous plan matrices were used to generate the query and reference of homogenous plans, as well as, the (269) SIB plan matrices, were used to generate the query and reference of SIB plans. For each OAR, a single merged reference matrix out of (181) homogenous plan matrices, and a single merged reference matrix out of (269) SIB matrices, were derived. The merged reference matrix was generated by grouping all voxels located at approximately same distance to PTV bin (bin size = 0.5 mm) from all reference plans into the same distance to PTV bin. The query set was

made up of the same plans (181) homogenous plan matrices and (269) SIB plan matrices, but this time each plan matrix was used separately without being merged (voxels at approximately same distance to PTV bin (bin size = 0.5 mm) in each query plan were grouped into same distance to PTV bin). The mean dose-at-distance function was calculated for the homogenous reference matrix and the SIB reference matrix, as well as, for each single query plan matrix. A schematic representation which summarizes the workflow is reported in Fig. 1.

2.3.3 Plan evaluation metrics: $M_{q,r}$ value

For both groups (homogenous and SIB VMAT plans), the query and reference dose distributions were compared by calculating the $M_{q,r}$ value. This metric value characterises the quality of OAR sparing of the query plan (index q) relative to the merged reference plans (index r) according to

$$M_{q,r} = \frac{1}{N_q} \sum_i n_{q,i} \cdot \frac{d_{q,i} - d_{r,i}}{\sqrt{\frac{s_{q,i}^2}{n_{q,i}} + \frac{s_{r,i}^2}{n_{r,i}}}} \quad . \tag{1}$$

Where $d_{q,i}$ is the mean dose at distance-to-PTV bin i of the query plan q of each OAR (bladder, rectum) and $d_{r,i}$ is the mean absorbed dose at distance-to-PTV bin i of the reference matrix r of each OAR. This mean dose difference indicates the shift between each query plan matrix compared to the reference matrix in each group for each OAR. Negative differences demonstrate that the OAR in the query plan is better spared., i.e., less mean dose in the query plan. $s_{q,i}^2$ and $s_{r,i}^2$ are the variances of the dose distributions at a given distance-to-PTV bin i for the query plan and the reference matrix. $n_{q,i}$ and $n_{r,i}$ are the numbers of voxels at distance-to-PTV bins i for the query plan and the reference matrix, respectively. This metric is derived based on a t-test as it compares the sample dose distribution over each single bin i in a query plan and the sample dose distribution in the matching bin in the reference matrix. This sample dose distribution difference over each individual bin is weighted by the number of voxels of the query plan at the same bin and summed over all the bins. After all, the summed

value is divided by the total number of voxels N_q of the query plan which is calculated as

$$N_q = \sum_i n_{q,i} \quad . \tag{2}$$

For every query plan of each treatment group (homogenous and SIB plans), $M_{q,r}$ of the rectum and $M_{q,r}$ of the bladder were calculated to investigate the obtained OAR dose sparing for the individual patient plan. $M_{q,r}$ values for both OARs represent the data points of this work.

2.3.4 Statistical Analysis

Mean values and standard deviations were calculated for $M_{q,r}$ of both OARs for each group (Table 3). Then, the confidence ellipses (50%, 80%, 85%, 90%, 95%) were drawn for the 2D normally distributed data points, $M_{q,bladder}$ versus $M_{q,rectum}$.

2.3.5 Model Validation

Patient plans between 80%-85%, 85%-90%, 90%-95% and 95%-100% probability ellipses to the upper right corner of longitudinal major axis (relatively high $M_{q,r}$ of both rectum and bladder) and to the lower right corner of longitudinal major axis (relatively high $M_{q,r}$ of bladder), were exported to TPS to be revised and re-optimized to potentially achieve better sparing for OARs while maintaining the best possible target coverage. [Fig. 1]. The re-planning was subjective based on the clinical experience of the oncologist and physicist in a trial and error process using MONACO® TPS. All the re-optimized plans were accepted for clinical treatment by an experienced physician. Afterwards, $M_{q,bladder}$ versus $M_{q,rectum}$. of the re-optimized plans were calculated and 2D normally distributed data points of re-optimized plans were added to the original probability ellipse bivariate normal distribution for each group. The numbers of improved plans after re-optimization to the total number of plans outside 80%, 85%, 90% and 95% probability ellipses were calculated.

Such quantification of successfully re-optimized plans is a step towards establishing a specific threshold above which re-examination and re-optimization of

the plans is strongly recommended, as they should have a relatively high probability of being improved.

The DVHs were calculated for the query plans outside 80%, 85%, 90% and 95% probability ellipses for the comparison of the plans before and after re-optimization.

2.3.6 Refinement of database

In the absence of a metric that incorporates patient anatomy to aid the acceptance of best optimum quality of treatment plans, sub-optimal plans still could be approved for patient treatment. To distinguish such sub-optimal plans, $M_{q,r}$ value was calculated for all the plans in the original data base. Plans located outside 80% probability ellipses to the right of longitudinal major axis were classified as sub-optimal and were replaced by the re-optimized plans to obtain an improved data base of optimal plans [Fig. 1].

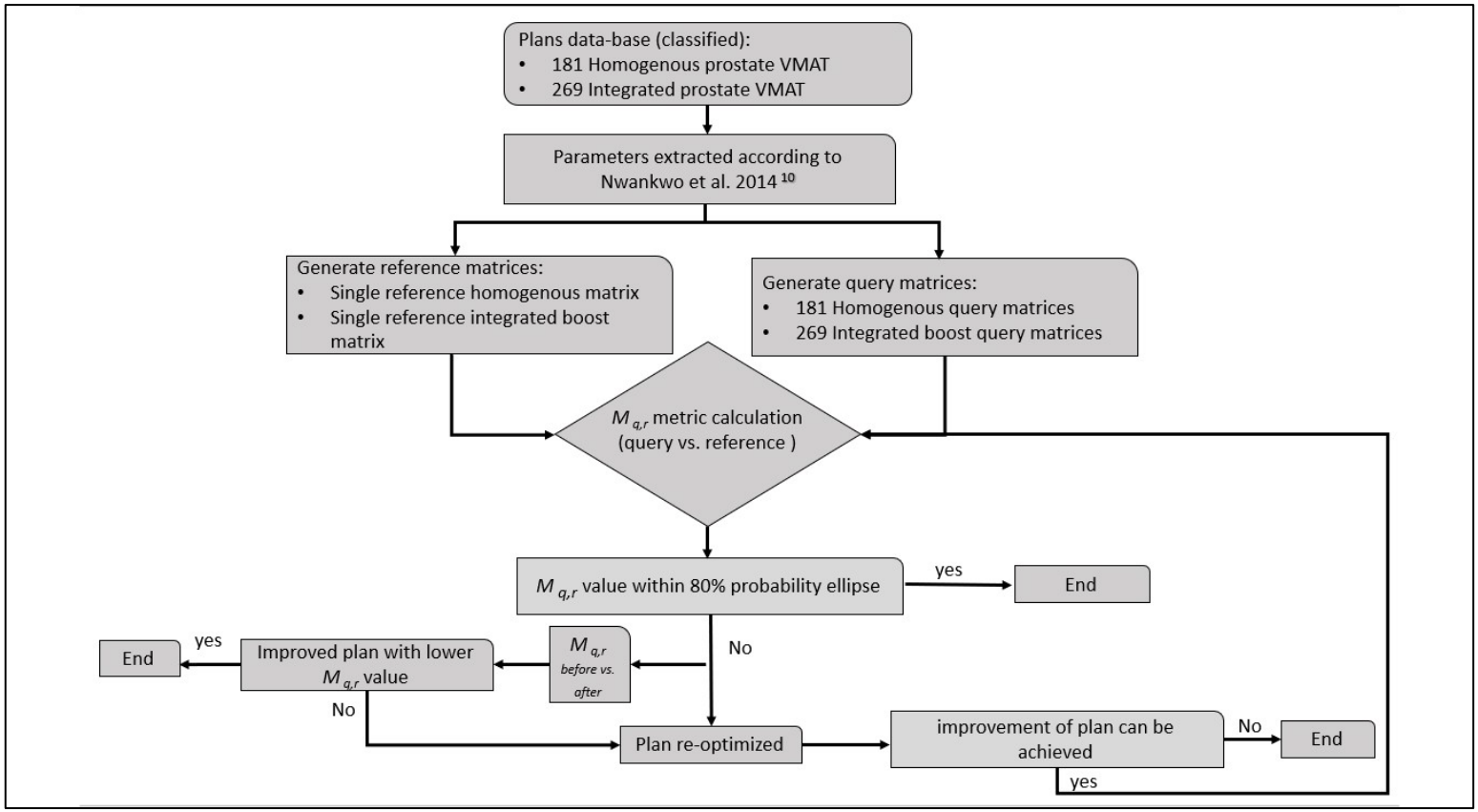

Figure 1. A schematic representation of the study workflow. Homogenous and SIB VMAT TPs form the data base of this study. The plans were used to generate both the reference merged matrix and the query matrices. For each group, a query matrix was compared with the single merged reference matrix by calculating the $M_{q,r}$ metric (Eq. (1)). If the $M_{q,r}$ value was between 80%- 85%, 85%- 90%, 90% - 95% and 95% - 100% probability ellipses, the plan was selected for re-optimization. After being re-optimized, the $M_{q,r}$ value of the resultant plan was calculated and checked whether or not it is improved (lower value is gained) . This process can be repeated until no further improvement of plans can be achieved.[37]

3 RESULTS

The results of the calculated $M_{q,r}$ values before and after re-optimization are presented in Figures 2 and 3.

The findings can be summarized as follows:

1. Homogenous plans: eight out of eleven plans (8/11 TPs) falling outside the 90% (containing four plans out of five (4/5 TPs) located outside the 95% ellipse) could be improved after re-planning to achieve better OARs sparing while preserving the same or better target coverage. Thus, 73% of the plans located outside the 90% probability ellipse were improved (Fig. 2(b)). Three out of four plans (3/4) falling outside the 80% (containing one out of two (1/2) located outside the 85% ellipse) were improved after re-planning (Fig. 2(c)).
2. SIB plans: six plans out of thirteen (6/13 TPs) located outside the 90% probability ellipse (with five plans out of eight (5/8 TPs) outside the 95% ellipse) were improved after re-planning which means 45% of the plans located out of the 90% were improved (Fig. 3(b)). One out of nine TPs (1/9) located outside the 80% (with one out of seven plans (1/7) falling outside the 85% ellipse) were improved after re-planning (Fig. 3(c)).

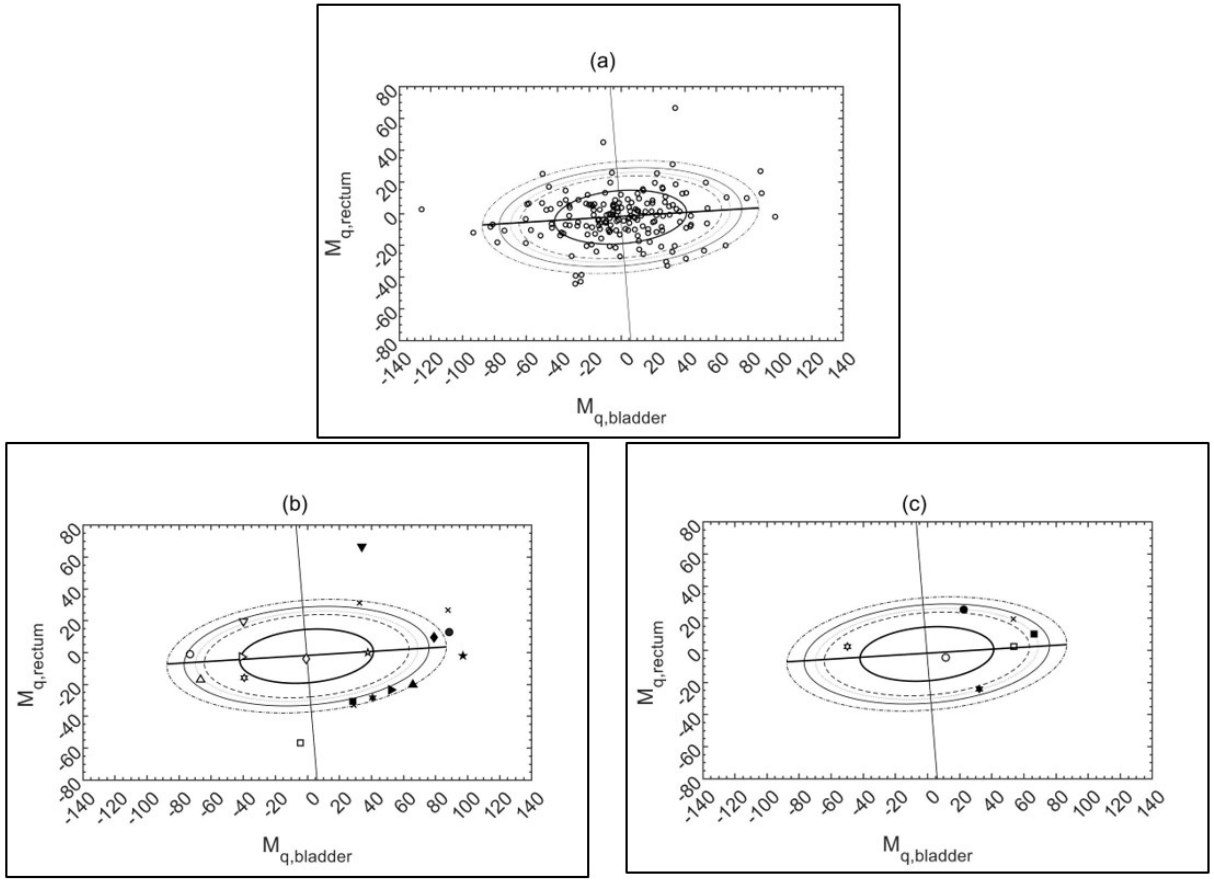

Figure 2(a) Probability ellipses of a bivariate normal distributed $M_{q,bladder}$ versus $M_{q,rectum}$ for homogenous VMAT prostate plans. Bold full line represents the 50% probability ellipse, dashed line, dotted line, full line and dashed-dotted line represent the 80%, 85%, 90% and 95% probability ellipses, respectively. The 80%, 85%, 95% and 90% probability ellipses were considered in this study. (b) TPs between the 90%-95% and 95%-100% probability ellipses before and after re-optimization. Each data points pair with the same symbol represent plans for the same patient. Filled symbols represent patient plans before re-optimization, while open symbols stand for the plans after re-optimization. Cross-shaped symbols represent patient plans which could not be improved by re-optimization. (c) TPs between the 80%-90% probability ellipses before and after re-optimization.[37]

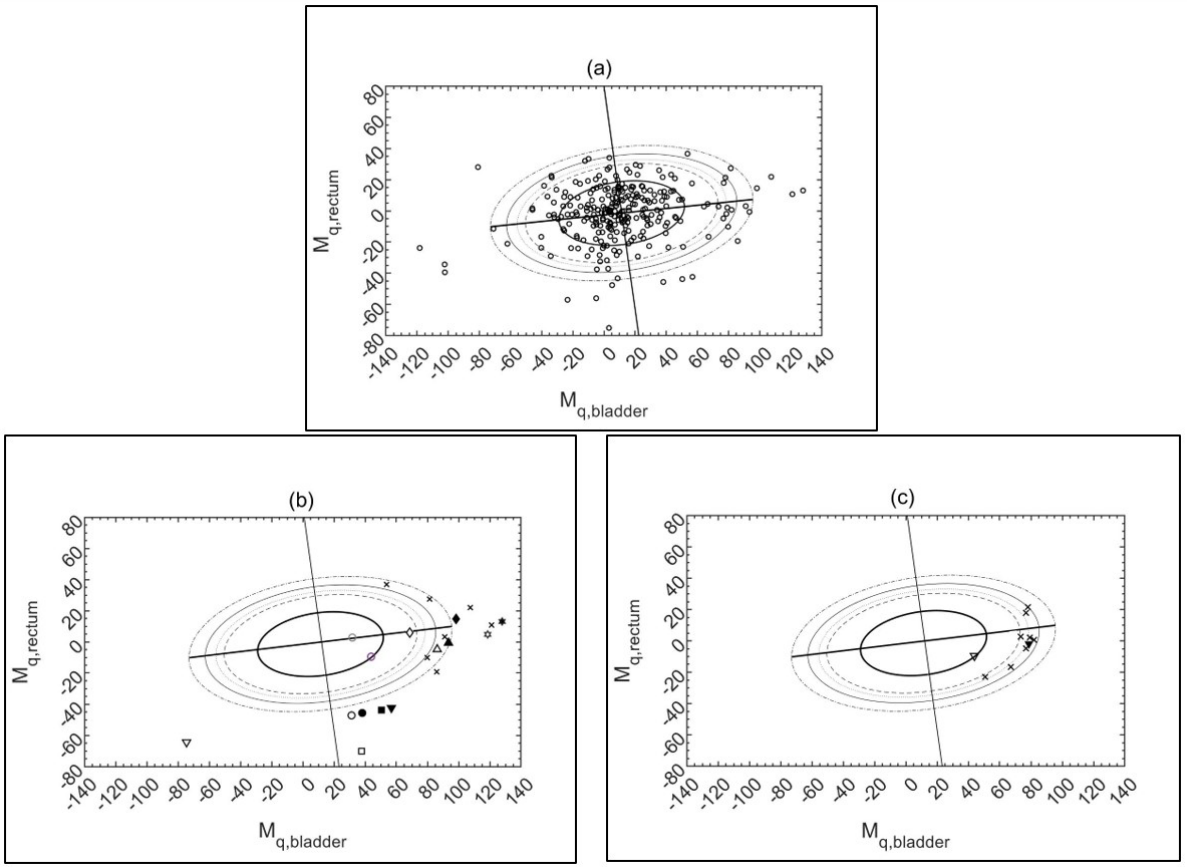

Figure 3 (a) Probability ellipses of a bivariate normal distributed $M_{q,bladder}$ versus $M_{q,rectum}$ for SIB VMAT prostate plans. Bold full line represents the 50% probability ellipse, dashed line, dotted line, full line and dashed-dotted line represent the 80%, 85%, 90% and 95% probability ellipses, respectively. The 80%, 85%, 95% and 90% probability ellipses were considered in this study. (b) TPs between the 90%-95% and 95%-100% probability ellipses before and after re-optimization. Each data points pair with the same symbol represent plans for the same patient. Filled symbols represent patient plans before re-optimization, while open symbols stand for the plans after re-optimization. Cross-shaped symbols represent patient plans which could not be improved by re-optimization. (c) TPs between the 80%-90% probability ellipses before and after re-optimization.[37]

Tables I and II show the $M_{q,r}$ values for the homogenous and SIB plans before and after re-optimization. In most of the plans, the quality of both rectum and bladder was improved with respect to the $M_{q,r}$ values. However, in the homogenous group of prostate plans, a relatively small increase of $M_{q,rectum}$ was observed [Table I], i.e., for the hexagram and right-sided triangle data points outside of the 90% ellipse and the pentagram data point out of the 95% ellipse [Figure. 2(b)]. But, this was within the clinically acceptable ranges and did not affect the obtained quality of the re-calculated plans.

Figures 4 and 5 show seven homogenous plans and four SIB plans, respectively, using their DVHs. Based on $M_{q,r}$ difference of the rectum of the homogenous plans (last column of table I), inverted triangle and triangle plans demonstrated the best and worst $M_{q,r}$ difference to the rectum outside 95%, respectively, while square and right-sided triangle plans demonstrated the best and worst $M_{q,r}$ difference to the rectum outside 90% respectively. Among the homogenous plans outside 80% probability ellipse, circle and hexagram plans had the best and worst $M_{q,r}$ difference to the rectum (last column of table I). However, for SIB plans (last column of table II), $M_{q,r}$ difference to the bladder indicates that inverted triangle plan has the best M_q difference while circle plan has the worst between plans before and after re-optimization outside 95%. Due to space limitations, only the DVH of the 13 plans in the two groups are shown.

Table 1. $M_{q,r}$ values of bladder and rectum, before and after re-optimization for homogenous VMAT prostate plans. Lower $M_{q,r}$ values indicate better quality, which implies better OAR sparing was achieved. Plans between 80%-85%, 85%-90%, 90%-95% and 95%-100% probability ellipses were considered.[37]

	Before re-optimization	After re-optimization	
data point shape	$M_{q,bladder}$	$M_{q,bladder}$	Difference
	95% -100%		
pentagram	97.0	37.7	59.3
circle	88.5	-73.2	161.7
inverted triangle	33.9	-39.9	73.8
triangle	65.8	-66.6	132.4
	90% - 95%		
square	28.3	-4.3	32.6
hexagram	40.7	-39.4	80.1
right-sided triangle	52.2	-40.9	93.1
diamond	79.0	-0.8	79.8
	85% - 90%		
square	66.4	29.9	36.5
	80% - 85%		
circle	22.5	11.3	11.2
hexagram	32.2	-49.9	82.1
	$M_{q,rectum}$	$M_{q,rectum}$	Differences
	95% -100%		
pentagram	-1.9	0.1	-2.0
circle	12.9	-1.0	13.9
inverted triangle	66.5	19.6	46.9
triangle	-20.2	-16.9	-3.3
	90% - 95%		
square	-30.5	-56.8	26.3
hexagram	-28.4	-15.8	-12.6
right-sided triangle	-23.4	-2.7	-20.7
diamond	9.7	-3.9	13.6
	85% - 90%		
square	10.2	2.3	7.9
	80% - 85%		
circle	25.4	-4.5	29.9
hexagram	-24.1	2.4	-26.5

Table 2. $M_{q,r}$ values of bladder and rectum, before and after re-optimization for SIB VMAT prostate plans. Lower $M_{q,r}$ values indicate better quality, which implies better OAR sparing was achieved. Plans between 80%-85%, 85%-90%, 90%-95% and 95%-100% probability ellipses were considered. Both rectum and bladder in all the SIB plans gained better sparing (lower $M_{q,r}$ values).[37]

	Before re-optimization	After re-optimization	
data point	$M_{q,bladder}$	$M_{q,bladder}$	Differences
	95% -100%		
circle	38.2	31.3	6.9
square	50.3	37.6	12.7
inverted triangle	56.9	-74.7	131.6
hexagram	127.8	118.6	9.2
diamond	98.2	68.6	29.6
	90% - 95%		
triangle	93.4	86.2	7.2
	85% - 90%		
inverted triangle	78.7	43.9	34.8
	$M_{q,rectum}$	$M_{q,rectum}$	Differences
	95% -100%		
circle	-45.7	-47.2	1.5
square	-43.8	-70.2	26.4
inverted triangle	-42.4	-64.4	22
hexagram	13.1	4.7	8.4
diamond	14.6	5.9	8.7
	90% - 95%		
triangle	-0.6	-4.8	4.2
	85% - 90%		
inverted triangle	-1.9	-9.6	7.7

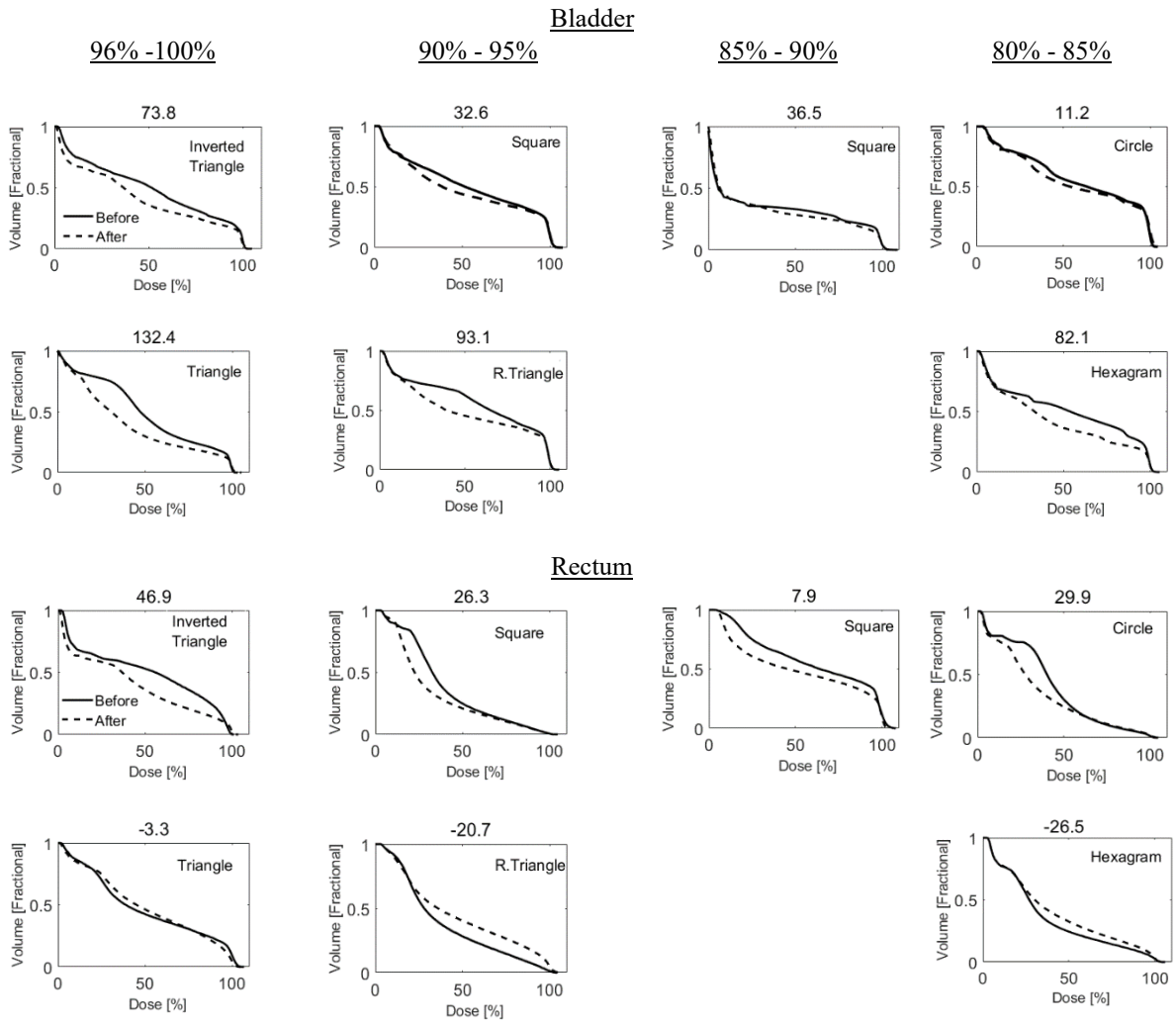

Figure 4. Validation of the model of the bladder and rectum of the homogenous plans. The DVHs were computed for both OARs of the query plans before and after re-optimization. Solid lines represent plans before re-optimization while dashed lines represent plans after re-optimization. Within each percentile range (column), the first two rows showed the best and worst plans based on the $M_{q,r}$ difference of plans (titles of subplots) before and after re-optimization.[37]

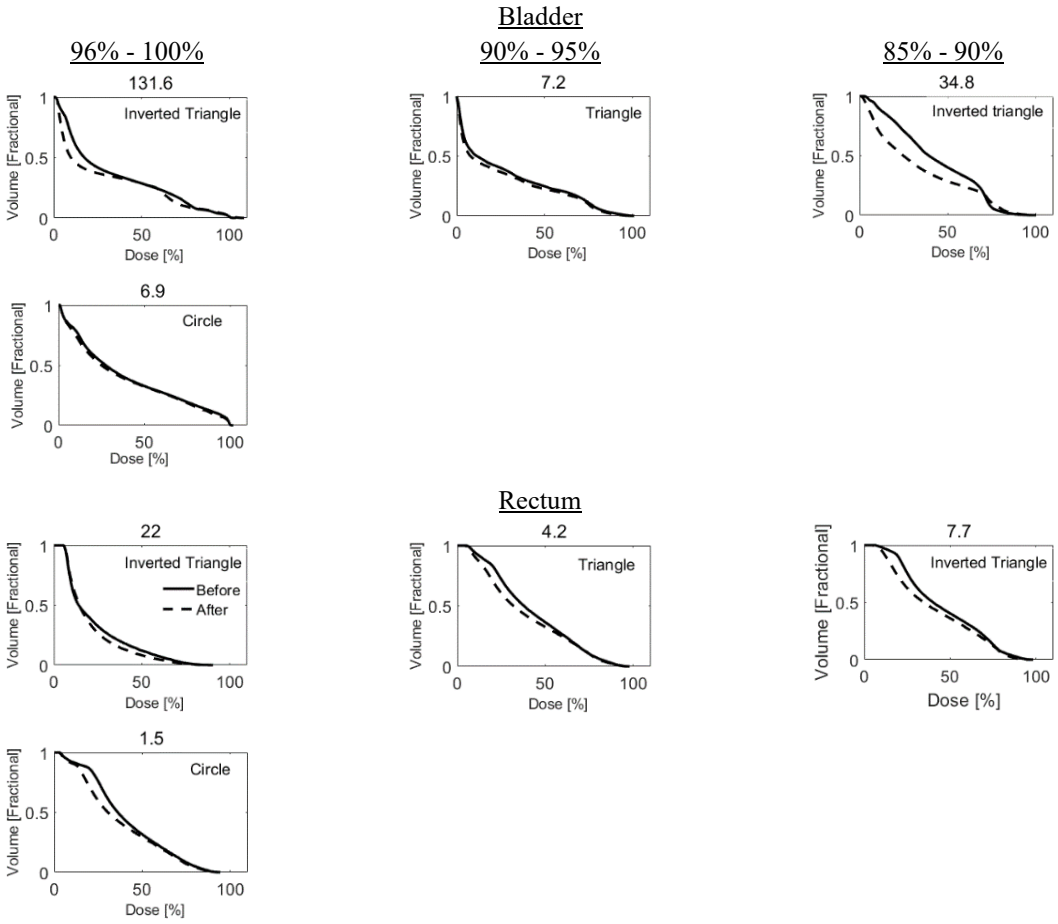

Figure 5. Validation of the model of the bladder and rectum of the SIB plans. The DVHs were computed for both OARs of the query plans before and after re-optimization. Within each percentile range (column), the first two rows showed the best and worst plans based on the $M_{q,r}$ difference of plans (titles of subplots) before and after re-optimization.[37]

Lastly, all the plans located outside 80% probability ellipse were replaced with the re-optimized plans. Thus, the reference set of both groups for both OARs were refined. Table III shows the $M_{q,r}$ values for the homogenous and the SIB plans before and after the data-base was refined.

Table 3. Summary of $M_{q,r}$ values - Mean ± Standard Error (SE), STD ± SE, the correlation parameter between bladder and rectum $M_{q,r}$ values (rho factor) - for both treatment groups before and after the data-base was refined. Lower $M_{q,r}$ values indicate better quality, which implies that better OAR sparing was achieved.[37]

		Homogenous TPs		SIB TPs	
Data-Base refined		Before	After	Before	After
$M_{q,rectum}$	Mean ± SE	- 2.2 ± 1.1	- 2.1 ± 1.0	-1.5 ±1.1	-1.5 ±1.1
	STD ± SE	14.6 ± 5.7	13.5 ± 0.7	17.8 ± 0.8	18.2 ± 0.8
$M_{q,bladder}$	Mean ± SE	-0.4 ± 2.6	-0.4 ± 2.4	11.4 ± 2.1	10.7 ± 2.0
	STD ± SE	35.5 ± 1.9	33.0 ± 1.7	34.5 ± 1.5	34.0 ± 1.4
rho factor		0.17	0.17	0.20	0.23

4 DISCUSSION

We developed and propose a quantitative KBRT algorithm to characterize the quality of radiotherapy treatment plans, which aids planners in identifying potential absorbed dose reductions for OARs during the treatment planning process. The algorithm was validated for the prostate but can be applied to other treatment sites. The first step is to determine the relationship between the positions of OARs (relative to the treatment volume) and their doses by analysing a reference database of prostate TPs accepted for patient treatment. This knowledge was used later to develop a QC quantitative approach that characterizes the quality of the plans to define a threshold that spots sub-optimal plans.

An objective decision on the quality of the radiotherapy TPs requires quantitative knowledge of what can be achieved for the particular anatomy.[23] Published dose-volume constraints used in evaluating plan quality are population-based and do not guarantee that the best possible plan is achieved for the varying patient anatomy. Our knowledge-based quantitative metric ensures that OAR doses are minimized to the extent permitted by the anatomy of the treated patient. The need to examine trade-offs at a fine level makes it desirable to model the absorbed dose at the voxel level. This is not addressed in most existing model-based dose prediction studies[23] since they handle summary dose metrics such as the DVH.

In this work, a knowledge-based algorithm was developed and evaluated to estimate a threshold for the achievable absorbed dose sparing in OARs to help radiotherapy planners to identify sub-optimal TPs. Each query plan was compared to the reference plans by calculating $M_{q,rectum}$ and $M_{q,bladder}$. (Eq. (1), p (7)). The $M_{q,r}$ metric was calculated based on the weighted differences of the mean doses at binned distances to the PTV surface. The 80%, 85%, 90% and 95% probability ellipses for the normal distributed data points (patient plans) of OARs were considered to define the threshold above which the plan is highly recommended to be re-optimized. The re-planning, although not the main focus of the study, was performed to confirm that plans falling below certain thresholds defined by this method could be significantly improved. The identified deficiencies in the sub-optimal plans fell in one of two categories; insufficient modulation level which leads to wider dose gradients than optimally possible, and insufficient use of cost-functions especially regarding the balance between bladder and

rectum that sometimes led to unnecessary dose spillage in one of the two organs. More specifically, this is due to inexperience and the absence of a tool that incorporates the anatomies of different patients to help the planner in the assessment of the plan quality. In addition, it can be due to the compromised sparing of an OAR for the benefit of another competing OAR which is deemed unsatisfactory focused.

The results of this work showed that the majority of the re-optimized plans were of greater quality when compared with the original plans in the data base for both SIB and homogenous VMAT prostate TPs especially outside 90%. Although much of plan quality assessment is subjective, and is best made by planners, eight out of eleven (8/11) homogenous prostate plans, as well as, six out of thirteen (6/13) SIB prostate plans outside the 90% probability ellipse had superior dose sparing of OARs with more uniform PTV coverage after re-optimization. In addition, three out of four (3/4) homogenous TPs and one out of nine (1/9) SIB TPs located outside 80% could be improved after re-planning. Thus, all the re-optimized plans would have potentially resulted in clinical improvements or clinically negligible differences when compared to the original plans. Re-optimizing the plans outside the 90% probability ellipse in both data groups led to more satisfying plans that better matched the planning dose volume objectives. The quantitative and qualitative ($M_{q,r}$ and DVHs) comparison of plans before and after re-optimized showed the efficiency and accuracy of this metric in evaluating the agreement between the compared plans.

The re-planning, although not the main focus of the study, was performed to confirm that plans falling below certain thresholds defined by this method could be significantly improved. It should be noted that, the observed dose reduction of OARs after the re-optimization could also have been achieved by modifying the planning constraints and objectives during optimization without guidance of a KBRT QC algorithm. However, the fact that the planners of the original plans did not achieve lower OARs doses and the original plans were approved, indicates the difficulty of finding the right set of constraints and objectives for planning and judging the quality of TP based on experience and intuition alone. If the planners had known that the OAR dose could have been decreased considerably without increasing the dose to the other OARs, it would have been unlikely that they would have approved the original plans. This emphasises the usefulness of a QC model for treatment planning. During the re-optimizing process, sometimes several iterations were performed to improve the

quality of plan but knowing what the achievable DVH should look like limited the number of iterations. Although, a longer planning time sometimes was required, it was within the same time range normally encountered to generate clinical acceptable plans. However, the achieved improvements in the sparing of OARs without degrading the coverage to the PTV, made the needed time and effort valuable since we know that the optimal plan has not been achieved based on prior knowledge.

The results of $M_{q,r}$ could be either positive or negative. It is important to note that the calculated $M_{q,r}$ metric was derived through other equations. For example, the squared weighted difference of mean dose at binned distance to PTV, M_2, was also calculated. But the calculated M_2 did not show the negative and positive sign of the results which affects the true meaning of the $M_{q,r}$ metric. Since the negative $M_{q,r}$ metric indicates that the query plan is below the reference, which means better sparing for the OARs. In the case of positive $M_{q,r}$ metric, the query plan is above the reference which means more weighted dose to the OARs is received in the query plan compared to reference plans. Another important point considered when comparing reference against query dose distribution is the trade-offs between OARs. In the planning process it is common to increase the sparing of one OAR at the expense of the other OAR. In our approach this is considered: For example, $M_{q,bladder}$ of a homogenous plan (Table I, pentagram) before re-optimization is 97, while for $M_{q,rectum}$ it is -1.9; this suggests that a prioritized sparing of rectum compared to bladder may be possible. Even if having negative scores of $M_{q,r}$ to both OARs indicating that both OARs are better spared compared to the reference plans, still one OAR should be lower than the other to indicate the priority of sparing in the competing OARs. An example for this is the homogenous plan (Table I, circle) after re-optimization with $M_{q,bladder}$ = -73.2 , $M_{q,rectum}$ = -1 which implies a prioritized sparing of bladder compared to rectum in the re-optimized plan. This adds a considerable worth to the calculated metric.

Also, this method showed that the chance to improve plans diminishes as we move inward toward the central part of ellipses. In homogenous plans (4/5 TPs) located outside 95% and (5/7 TPs) between 90% and 95% were able to be re-optimized. In SIB (5/8 TPs) outside 95%, as well as, (1/5 TPs) between 90% and 95% achieved better OARs sparing when re-optimized. However, it is still suggestive to check and improve the quality of the plans within 80%, 50%...etc. In fact, it could be argued that with sufficient time and experience, most TPs could be improved. Although this is true,

a cost-benefit consideration must be made: Is it worth investing hours to improve a plan to a level (relative to a reference) that is clinically irrelevant? As the gains are the smaller the more we move towards the centre of the ellipses, a threshold ellipse must be defined outside of which the plans are to be improved. It is also important to note that the use of clinically accepted plans which have relatively high quality is quite important because the algorithm is based on the mean values of dosimetric features of the used reference plans. We believe that this method can be used to eliminate sub-optimal plans iteratively from the reference set (removing plans with high positive $M_{q,r}$ values from the reference set while keeping those with negative or low $M_{q,r}$ value) will strict the characterizing capability of the used method and increase its efficiency. Therefore, the reference set of both groups for both OARs were refined by replacing all the plans outside 80% probability ellipse with the matching re-optimized plans with improved quality

Several studies have demonstrated the importance of adoption of QC methods to predict and quantify the achievable OAR sparing to provide planning consistency.[15,16] Wu et al.[15] proposed a QC method to optimize the DVHs of the OARs in new TPs based on an OVH descriptor to identify related patients. They reported a clinically significant excess radiation dose that was delivered to patient as a result of insufficient plan QC. Other recent advances in knowledge-based methods used machine learning[12, 38]. Zhu et al.[21,22] proposed a quantitative evaluation tool based on machine learning to characterize the relation between DVH and anatomical features, which could be used later to optimize plan quality.[23,24] Although these approaches help planners to achieve better OARs sparing based on the DVH-guidance. The inter-patient anatomical variations and their impact on the OARs sparing need more than DVH objectives to achieve lowest possible dose to OARs.

An interesting approach that has been investigated to overcome these issues is voxel-based dose prediction. Nwankwo et al.[32] proposed a KB QC approach depending on the mean dose difference at voxel level between the compared plans. Although, in the proposed approaches of this work and Nwankwo et al [32], the same KB algorithm was used, our approach is unique as it introduces a threshold to spot the sub-optimal plans that are highly recommended to be re-optimized using $M_{q,r}$ metric. In addition to the mean dose difference at binned distance to PTV, other factors that expected to have an impact on the $M_{q,r}$ metric were considered, to calculate a more realistic outcomes.

For example, the number of estimated standard deviations of the sample mean dose from it is expected value were considered. This estimation includes different variables that directly affect the $M_{q,r}$ metric values. Firstly, the variances of the dose distributions at a given distance-to-PTV bin for the query plan and the reference matrix that give a feedback on the dose scattering around the mean over the compared bins. Secondly, the numbers of voxels at distance-to-PTV bins for both query and reference, which corresponds to the number of observations, since this metric was derived in analogy to a t-test.[39] If we have a large number of observations and all of these observations are close to the sample mean (large $\boldsymbol{n}$, small $\boldsymbol{s}$), we can be confident that our estimate of the sample dose distribution is fairly accurate, which results in a small $M_{q,r}$ value. The large number of plans used in this approach database enables the approach making more accurate calculations, since it allows more observed variations in organ geometries.[40] However, a limitation of this approach is that the quality of the new plan strongly depends on the quality of the reference plans.

Finally, this KBRT QC approach is expected to be efficient even with more complicated treatment sites and their corresponding treatment techniques, i.e., head and neck tumours and Hyper-Arch technique. Even-though these techniques improve the sparing of the adjacent critical tissue while delivering a more conformal dose to the PTV, but they still need to fulfil the DVH-guidance. However, judging the real benefits of integrating this algorithm for more complex treatment sites / techniques is a subject that requires further investigations. We believe that the adoption of this QC method in the planning process will provide better planning efficiency in radiation therapy centres, especially ones with a lower experience level.

5 SUMMARY

Inverse-treatment planning has a remarkable capability of “dose painting” to minimize the dose to OARs while achieving prescription coverage to PTV. However, the task of creating/optimizing a “best plan” continues to be difficult and time-consuming. The anatomical variation of multiple OARs, their distant location to PTV create types of trade-offs such that the quality of a plan and the speed of planning depend heavily on the experience of physicians and planners. Therefore, despite the computational developments and the years of experience in plan optimization, inverse treatment planning is still a time-consuming process depending on planners’ subjective decisions. To tackle these challenges, a number of approaches based on KBRT have been successfully developed, i.e. KB QC methods to predict OARs sparing. Furthermore, most of the efforts to standardize and automatize inverse treatment planning are DVH-based KB approaches. However, these methods cannot provide a reasonable estimation of the achievable patient-specific dose distributions since they lack the objective standards to evaluate the quality of the plans depending on the anatomical variations.

We address this problem by developing a knowledge-based QC algorithm of using clinically approved prostate VMAT TPs, whereby we used the prior knowledge of treatment planning to achieve the greatest possible rectum-bladder sparing without compromising target coverage. This QC approach aims to characterize the quality of the TPs by evaluating the doses of the OARs (rectum and bladder) in order to define a threshold that detects sub-optimal treatment plans based on the variability of patient anatomy.

Therefore, rather than predicting DVHs, this quantitative metric ensures that OAR doses are minimized to the extent permitted by the anatomy of the individual treated patient.

The knowledge-based database consisted of 450 VMAT prostate plans that were divided into 181 homogenous prostate plans and 269 SIB prostate plans. For both of the planning groups, the reference sets matrix and the query set matrices were generated for each OAR. For the step of plan quality evaluation, $M_{q,r}$ metric was developed in analogy to a t-test. More specifically, this metric compares the weighted differences of the mean doses at binned distances to the PTV surface of the individual

query plan matrix against the reference matrix. The 90% probability ellipse of the 2 D normally distributed data points ($M_{q,\ bladder}$ *versus* $M_{q,\ rectum}$) were considered to define a threshold above which the treatment plan is strongly recommended to be re-optimized. Afterward, for both OARs, the DVH was calculated to compare the plans before and after being re-optimized for the purpose of validating the approach.

The results of this work demonstrated that the majority of the re-optimized plans had superior quality compared to the original plans in the database for both groups, especially outside 90%. More precisely, $M_{q,r}$ values of bladder and rectum demonstrated lower values after TPs were reoptimized for both groups (homogenous and SIB). 8/11 of the homogenous plans in addition to 6/13 of the SIB plans outside the 90% probability ellipse were of greater quality with respect to OARs dose sparing while achieving a better target coverage. Also, 3/4 of the homogenous TPs and 1/9 of the SIB TPs between 80% - 90% were improved. The quantitative and qualitative ($M_{q,r,}$ and DVHs) comparison of plans before and after re-optimized confirms the efficiency of this metric for investigating the quality of plans.

In general, the patient-specific objective measures of plan quality are not currently widely used in clinical practice. To improve treatment plan evaluation, an objective plan quality assurance (QA) tool is needed. If this QC method of using a knowledge base of treatment plans can be made available for use at institutions, it could lead to more consistent quality of treatment planning across institutions.

6 REFERENCES

1. Litwin MS. Quality of life following definitive therapy for localized prostate cancer: potential impact of multiple therapies. *Current opinion in urology.* 2003;13(2):153-156.
2. Bucci MK, Bevan A, Roach M, 3rd. Advances in radiation therapy: conventional to 3D, to IMRT, to 4D, and beyond. *CA: a cancer journal for clinicians.* 2005;55(2):117-134.
3. Bortfeld T, Thieke C. Optimization of Treatment Plans, Inverse Planning. In: Schlegel W, Bortfeld T, Grosu A-L, eds. *New Technologies in Radiation Oncology.* Berlin, Heidelberg: Springer Berlin Heidelberg; 2006:207-220.
4. Dursun P, Taşkın ZC, Altınel İK. The determination of optimal treatment plans for Volumetric Modulated Arc Therapy (VMAT). *European Journal of Operational Research.* 2019;272(1):372-388.
5. Mohan R, Wang X, Jackson A, et al. The potential and limitations of the inverse radiotherapy technique. *Radiotherapy and Oncology.* 1994;32(3):232-248.
6. Lawrence TS, Kessler ML, Ten Haken RK. *Clinical Interpretation of Dose-Volume Histograms: The Basis for Normal Tissue Preservation and Tumor Dose Escalation.*
7. Otto K. Volumetric modulated arc therapy: IMRT in a single gantry arc. *Medical physics.* 2008;35(1):310-317.
8. Kopp RW, Duff M, Catalfamo F, Shah D, Rajecki M, Ahmad K. VMAT vs. 7-field-IMRT: assessing the dosimetric parameters of prostate cancer treatment with a 292-patient sample. *Medical dosimetry : official journal of the American Association of Medical Dosimetrists.* 2011;36(4):365-372.
9. Thieke C, Bortfeld T, Niemierko A, Nill S. From physical dose constraints to equivalent uniform dose constraints in inverse radiotherapy planning. *Medical physics.* 2003;30(9):2332-2339.
10. Michalski JM, Moughan J, Purdy J, et al. Effect of Standard vs Dose-Escalated Radiation Therapy for Patients With Intermediate-Risk Prostate Cancer: The NRG Oncology RTOG 0126 Randomized Clinical Trial. *JAMA Oncology.* 2018;4(6):e180039-e180039.
11. Teoh M, Clark CH, Wood K, Whitaker S, Nisbet A. Volumetric modulated arc therapy: a review of current literature and clinical use in practice. *The British Journal of Radiology.* 2011;84(1007):967-996.

12. Peng F, Jia X, Gu X, Epelman MA, Romeijn HE, Jiang SB. A new column-generation-based algorithm for VMAT treatment plan optimization. *Phys Med Biol.* 2012;57(14):4569-4588.

13. Ghandour S, Matzinger O, Pachoud M. Volumetric-modulated arc therapy planning using multicriteria optimization for localized prostate cancer. *Journal of applied clinical medical physics / American College of Medical Physics.* 2015;16:5410.

14. Good D, Lo J, Lee WR, Wu QJ, Yin FF, Das SK. A knowledge-based approach to improving and homogenizing intensity modulated radiation therapy planning quality among treatment centers: an example application to prostate cancer planning. *Int J Radiat Oncol Biol Phys.* 2013;87(1):176-181.

15. Wu B, Ricchetti F, Sanguineti G, et al. Patient geometry-driven information retrieval for IMRT treatment plan quality control. *Med Phys.* 2009;36(12):5497-5505.

16. 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(2):545-551.

17. Wu B, Ricchetti F, Sanguineti G, et al. Data-driven approach to generating achievable dose-volume histogram objectives in intensity-modulated radiotherapy planning. *Int J Radiat Oncol Biol Phys.* 2011;79(4):1241-1247.

18. Chanyavanich V, Das SK, Lee WR, Lo JY. Knowledge-based IMRT treatment planning for prostate cancer. *Med Phys.* 2011;38(5):2515-2522.

19. Wang Y, Zolnay A, Incrocci L, et al. A quality control model that uses PTV-rectal distances to predict the lowest achievable rectum dose, improves IMRT planning for patients with prostate cancer. *Radiother Oncol.* 2013;107(3):352-357.

20. Petit SF, Wu B, Kazhdan M, et al. Increased organ sparing using shape-based treatment plan optimization for intensity modulated radiation therapy of pancreatic adenocarcinoma. *Radiother Oncol.* 2012;102(1):38-44.

21. 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(11):6868-6878.

22. 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(2):719-726.

23. 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(12):7446-7461.

24. Schubert C, Waletzko O, Weiss C, et al. Intercenter validation of a knowledge based model for automated planning of volumetric modulated arc therapy for prostate cancer. The experience of the German RapidPlan Consortium. *PLoS One.* 2017;12(5):e0178034.

25. Tol JP, Delaney AR, Dahele M, Slotman BJ, Verbakel WF. Evaluation of a knowledge-based planning solution for head and neck cancer. *Int J Radiat Oncol Biol Phys.* 2015;91(3):612-620.

26. Zieminski S, Khandekar M, Wang Y. Assessment of multi-criteria optimization (MCO) for volumetric modulated arc therapy (VMAT) in hippocampal avoidance whole brain radiation therapy (HA-WBRT). *J Appl Clin Med Phys.* 2018;19(2):184-190.

27. Craft DL, Halabi TF, Shih HA, Bortfeld TR. Approximating convex pareto surfaces in multiobjective radiotherapy planning. *Med Phys.* 2006;33(9):3399-3407.

28. Hussein M, South CP, Barry MA, et al. Clinical validation and benchmarking of knowledge-based IMRT and VMAT treatment planning in pelvic anatomy. *Radiother Oncol.* 2016;120(3):473-479.

29. Monz M, Kufer KH, Bortfeld TR, Thieke C. Pareto navigation: algorithmic foundation of interactive multi-criteria IMRT planning. *Phys Med Biol.* 2008;53(4):985-998.

30. Teichert K, Suss P, Serna JI, Monz M, Kufer KH, Thieke C. Comparative analysis of Pareto surfaces in multi-criteria IMRT planning. *Phys Med Biol.* 2011;56(12):3669-3684.

31. Shiraishi S, Tan J, Olsen L, Moore K. Knowledge-based prediction of plan quality metrics in intracranial stereotactic radiosurgery. *Med Phys.* 2015;42:908.

32. Nwankwo O, Sihono DS, Schneider F, Wenz F. A global quality assurance system for personalized radiation therapy treatment planning for the prostate (or other sites). *Phys Med Biol.* 2014;59(18):5575-5591.

33. Nwankwo O, Mekdash H, Sihono DS, Wenz F, Glatting G. Knowledge-based radiation therapy (KBRT) treatment planning versus planning by experts: validation of a KBRT algorithm for prostate cancer treatment planning. *Radiat Oncol.* 2015;10:111.

34. Deasy JO, Blanco AI, Clark VH. CERR: a computational environment for radiotherapy research. *Med Phys.* 2003;30(5):979-985.

35. Brown WE, Sung K, Aleman DM, Moreno-Centeno E, Purdie TG, McIntosh CJ. Guided undersampling classification for automated radiation therapy quality assurance of prostate cancer treatment. *Med Phys.* 2018;45(4):1306-1316.

36. Stanhope C, Wu QJ, Yuan L, et al. Utilizing knowledge from prior plans in the evaluation of quality assurance. *Phys Med Biol.* 2015;60(12):4873-4891.

37. Alnaalwa B, Nwankwo O, Abo-Madyan Y, Giordano FA, Wenz F, Glatting G. A. knowledge-based quantitative approach to characterize treatment plan quality: Application to prostate VMAT planning. *Med Phys.* 2020;48(1):94-104

38. Babier A, Mahmood R, McNiven AL, Diamant A, Chan TCY. Knowledge-based automated planning with three-dimensional generative adversarial networks. *Med Phys.* 2020;47(2):297-306.

39. Kim TK. T test as a parametric statistic. *Korean J Anesthesiol.* 2015;68(6):540-546.

40. Boutilier JJ, Craig T, Sharpe MB, Chan TC. Sample size requirements for knowledge-based treatment planning. *Med Phys.* 2016;43(3):1212-1221.

7 CURRICULUM VITAE

PESONAL INFORMATION

First Name and Family Name:	Buthayna Alnaalwa
Date of birth:	02 Nov 1985
Place of birth:	Tulkarm, Palestine

2016 – Present	PhD student in Medical Physics Medical Faculty Mannheim Heidelberg University, Germany.
2009 – 2010	M.Sc. Medical Physics-radiotherapy, Mannheim Medical Faculty, Heidelberg University, Germany.
2004 – 2008	B.Sc. Radiologic Technology program, Faculty of Allied Medical Science, Jordan University of Science and Technology, Jordan
2003 – 2004	General Secondary Certificate, Al-Adawiah Sec. Girls School, Tolkarm, Palestine.

PROFESSIONAL EXPERIENCE

April/2017-Present	PhD student, A knowledge-based approach to improve treatment planning quality: application to Prostate VMAT has been developed. Mannheim Medical Faculty, Heidelberg University.
April/2011 – Sep/2012 & Sep/2013– Dec/2013	Lecturer, College of Allied Medical Sciences, Arab American University, Jenin-Palestine. (Field: Medical Imaging).
April /2014 – Aug/2016	Lecturer, College of Medicine and Health Sciences, An-Najah National University, Nablus-Palestine. (Field: Medical Physics).

PUBLICATIONS

1. Alnaalwa B, Nwankwo O, Abo-Madyan Y, Giordano FA, Wenz F, Glatting G. A. knowledge-based quantitative approach to characterize treatment plan quality: Application to prostate VMAT planning. *Med Phys.* 2020;48(1):94-104

2. Alnaalwa B, Obioma N, Yasser AM, Frank GA, Frederik W, Gerhard G. A knowledge-based quantitative approach to characterize treatment plan quality: application to prostate VMAT planning. *DEGRO (2020)* Germany

WORKSHOPS ATTENDANCE

June/2010	Biomedical Engineering / Computational Bio-photonics Workshop, Mannheim-Germany.
August/2009	Guided and Robotic Radiotherapy Workshop**,** Mannheim-Germany.
April /2007	The First Arab Radiology Congress, Dead Sea, Jordan.

AWARDS

Aug/2009 - Nov/2010	Deuscher Akademischer Austauschdienst (DAAD) scholarship.
March/2010-July/2010	Mannheim Medical Faculty award to write master thesis in USA, Oregon Health & Science University.
April/2017-March/2020	Islamic Development Bank (IDB) Merit scholarship programme

FURTHER KNOWLEDGE

Computer Programming	MATLAB
Languages	Arabic (mother language) English (TOEFL) German Language (B1)

8 ACKNOWLEDGEMENT

First and foremost, praises and thanks to the God, the Almighty, for His showers of blessings throughout my research work to complete the research successfully.

I wish to express my sincere gratitude to Prof. Dr. rer. nat. Gerhard Glatting, for giving me the opportunity to work in his group and providing invaluable guidance throughout this research. It was a great privilege and honour to work and study under his guidance.

I am very grateful to Dr. Yasser Abo-Madyan for the useful discussions that I had with him. His vision, sincerity and motivation have deeply inspired me. His guidance helped me in all the time of research and writing of this thesis. I would like to thank Dr. Obioma Nwankwo, who started this project and gave me the initial introduction.

I am extremely grateful to my mother for her love, prayers, caring and sacrifices for educating and preparing me for my future. I am very much thankful to my husband and my kids for their love, understanding, prayers and continuing support to complete this research work.

Also, I would like to thank the scholarship of the Islamic Development Bank-Merit Scholarship Program (funding code 600031512) for the full funding support for this research study.

Finally, my thanks go to all the people who have supported me to complete the research work directly or indirectly.