SAPIENZA
UNIVERSITÀ DI ROMA

MASTER DI SECONDO LIVELLO IN “CALCOLO SCIENTIFICO”

A.A.2016/2017

APPLICATIONS OF RADIOMICS TECHNIQUES TO CLINICAL IMAGES AND CLINICAL DATA OF CANCER PATIENTS

Candidata

Smeralda Sparacino

Tutor Esterno

Dott.ssa Lidia Strigari

Tutor Interno

Prof. Camillo Cammarota

Index

Chapter 1

Introduction

Radiation therapy is therapy using ionization radiation, generally as part of cancer treatment to control or kill alignment cells and normally delivered by a linear accelerator. The aim of radiotherapy is to achieve a high probability of local tumor control (tumor control probability, TCP) at a low risk of normal tissue complications (normal tissue complication probability, NTCP).

In radiation therapy, treatment planning is the process in which a team consisting of radiation oncologists, radiation therapist, medical physicist and medical dosimetrist plan the appropriate external beam radiotherapy or internal brachytherapy treatment technique is optimized and customized for a cancer patient.

The work investigates the dose distribution obtained using several radiotherapy techniques, which are 3D conformal radiation therapy (3D-CRT) with/without Wedge or Field in Field techniques (FiF), the intensity-modulated radiation therapy (IMRT), the Volumetric Arc Therapy (VMAT). These techniques were applied for three representative patients with prostate cancer.

Radiomics is a field of medical study that aims at extracting a large amount of quantitative features from medical images using data characterization algorithms. These features, termed radiomic features, have the potential to uncover disease characteristics that fail to be appreciated by the naked eye.

In this work, radiomics has been used to identify more appropriate indicator able judging the more appropriate technique for the prostate cancer patients able to increase the TCP while sparing the organs at risk (bladder and rectum). Several techniques such as; IMRT, VMAT, Wedge field, Box and Field-in-Field were tested by calculating the radiomics features based on the dose distribution. The tumor control probability (TCP) and normal tissue complication probability (NTCP) have been calculated for each dose distribution to be used as gold standard for selecting the optimal plans and assess a planning score between rival plans.

Chapter 2

Materials and methods

2.1 Image/based treatment

Treatment is planned by using 3-D computed tomography (CT) or magnetic resonance (MRI) images of the patient in conjunction with computerized dose calculations to determine the dose intensity pattern that will best conform to the tumor shape.

2.2 Definition of Contouring, CTV, PTV and OAR

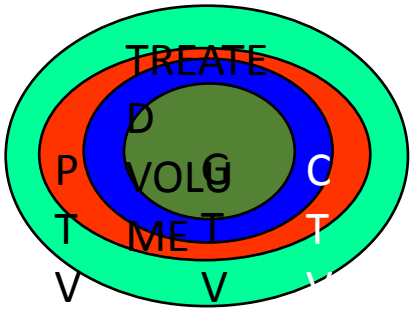
Target volume definition (Contouring) requires an assessment of clinical information and imaging to know the location of the primary tumor. The site known to have cancer and the areas at risk of tumor spread and defined on a series of cross-sectional images on the computer to create a 3-dimentional volume (3D) of part of the body to which a treatment dose of radiotherapy can be applied.

The GTV (Gross Tumor Volume) is essentially the gross demonstrable, palpable or visible or demonstrable malignant growth. It is the easiest volume to define, though not necessarily to localize. It is what can be seen, measured or imaged as well as a primary site. Typically, GTV corresponds to the part of the tumor where the tumor cell density is highest.

This may have implications for choice of radiotherapy dose, since tumor control requires a higher dose if the initial tumor cell number is larger. If the tumor has been removed, GTV cannot be identified.

The CTV (Clinical Target Volume) contains the demonstrable GTV plus a margin for sub-clinical disease spread. There may be different CTV, for example in a regional lymph node, where there is no obvious GTV present. The CTV is important because this volume must be adequately treated if cure is to be achieved. It is assumed that the tumor cell density in the CTV is lower than in the GTV, and consequently the radiotherapy dose may be lower.

The PTV (Planning Target Volume) is a geometric concept designed to ensure that the radiotherapy prescription dose is the dose delivered to the CTV. It takes in account the net effect of all possible geometrical inaccuracies and it is related to the **isocenter** of the linear accelerator rather than to the anatomy of the patient.



Organ at risk (OAR); (ICRU report 50, 62 and 83), also known as a critical structures, are anatomical structures with important functional properties located in the vicinity of the target volume. They have to be considered in

the treatment planning, since irradiation can cause pathological changes in the normal tissue, with irreversible functional consequences. In the prostate case organ at risk are rectum and bladder.

2.3 Radiation therapy treatment plans

2.3.1 Three dimensional conformal radiation therapy

The first radiotherapy technique is the 3D conformal radiation therapy (3D CRT); a cancer treatment that shapes the radiation beams to match the shape of the tumor, while sparing OARs by using MLC.

3D conformal therapy is not only to shape the field around the PTV, is to shape (paint, conform …) the highest possible absorbed dose around the PTV, lowering the dose in OARs, during the single fraction and for all the fractions of the treatment schedule.

Conformal radiation therapy: a procedure that uses computers to create a 3-dimensional picture of the tumor in order to target the tumor as accurately as possible and give it the highest possible dose of radiation while sparing normal tissue as much as possible. It is also known as 3-D or conformational radiation therapy.

2.3.1.1 Forward planning procedure

The standard method used in radiotherapy is to address the prescribed dose to PTV from multiple beams, in order to "concentrate" the

prescribed dose mainly in the PTV from different directions, and "spread" the unwanted dose in different areas in the bodies, with the organs at risk partially or completely avoided from some or all beams. Occasionally, wedges or compensators are used to modify the intensity profile to offset contour irregularities and/or produce more uniform composite dose distributions. The plan optimization is based on the selection of the weight of the each conformal beam in order to guarantee the homogeneity in dose distribution within the target and avoiding the over dosage in the surrounding tissues. A trial and see process is adopted to find the "best" plan in the forward planning procedure. At the end of optimization the planner identify the solution with highest achievable performance under the given constraints, maximizing the target coverage and minimizing the undesired dosage.

2.3.1.2 Wedge and Field- in -Field Technique

Wedges are commonly used as beam modifying devices in radiation therapy to optimize the target volume dose distribution. In this context, a variety of wedge filters is available with modern Linear accelerators. Varian's CLINAC 2100C provides two types of wedges the physical wedges (i.e. 15^{o}, 30^{o}, 45^{o} and 60^{o}) and enhancing dynamic wedges (i.e. 10^{o} 15^{o}, 20^{o}, 25^{o}, 30^{o}, 45^{o} and 60^{o}). The Physical Wedges (PW) can be inserted in the treatment head of linear accelerator in four different orientations (in, out, left and right). In the case of Enhanced Dynamic Wedge (EDW), the required dose distribution can be achieved by one of the collimator jaws motion in two different directions (in and out).

The Field-in-Field (FiF) technique is an often-used alternative to wedged fields in tangential irradiation for the treatment of prostate cancer. A simplified FiF technique that planers can easily achieve and which improves dose uniformity in the prostate volume is used in this work.

2.3.2 Inverse planning procedure

In modern radiation therapy departments, most treatments are delivered with radiation beams having a non-uniform intensity across the field named intensity-modulated radiation therapy (IMRT) or volumetric arc therapy (VMAT).

2.3.2.1 IMRT

IMRT is a radiation therapy technique in which a no uniform fluence is delivered to the patient from any given position of the treatment beam to optimize the composite dose distribution. The process for automatically obtaining an IMRT is called intensity modulation optimization. IMRT is as an advanced modality for delivering a high-precision radiotherapy in the specific areas within the tumor (or malignant cells) by using linear accelerators. It is a delivery technique that allows practically unlimited control over shaping of dose distribution to fit tumors of complex shape while sparing critical normal tissues in close proximity. The planner specifies the treatment criteria for plan optimization, while the "inverse planning" determines the optimal fluence profiles for a given set of beam directions. The generated fluence files are electronically transmitted to the linear accelerator, which is computer controlled, that is, equipped with the

required software and hardware to deliver the intensity-modulated beams as calculated. The treatment-planning program divides each beam into a large number of beamlets and determines optimum setting of their fluences or weights. The clinical implementation of IMRT requires at least two systems, a treatment-planning computer system and a MLC, i.e. a system of delivering the no uniform fluences as planned.

IMRT is being used most extensively to treat cancers of the prostate cancer, in fact, of all the sites, the prostate gland has received the greatest attention because of the greater degree of dose conformity that can be achieved compared to the conventional techniques, including 3-D conformal. Dose conformity is a "double-edged sword" with more normal tissue sparing on the one hand and greater possibility of target miss on the other. It is important to remember that an image-based treatment plan cannot fully account for the true extent of clinical target volume, the accurately applicable TCP and NTCP, and the natural motion of target volume and organs at risk. Too much emphasis on dose conformity can backfire because of these unavoidable uncertainties, resulting in inadequate target coverage for example.

Typically, combinations of multiple intensity-modulated fields coming from different beam directions produce a custom tailored radiation dose that maximizes tumor dose while also minimizing the dose to adjacent normal tissues. Because the ratio of normal tissue dose to tumor dose is reduced to a minimum with the IMRT approach, higher and more

effective radiation doses can safely be delivered to tumors with fewer side effects compared with conventional radiotherapy techniques. IMRT also has the potential to reduce treatment toxicity, even when doses are not increased.

The basic difference with the conventional radiotherapy (including 3-D CRT) is that IMRT provides an extra degree of freedom, that is, intensity modulation, in achieving dose conformity. Especially targets of concave shape surrounding sensitive structures can be treated conformly with steep dose gradients outside the target boundaries; a task that is almost impossible to accomplish with conventional techniques. For localized lesions in any part of the body, IMRT compares well with or exceeds the capabilities of other techniques or modalities. Additionally, IMRT is not limited by target size or its location. Like intensity-modulated radiation therapy, 3D conformal therapy targets cancer while sparing healthy tissue. However, IMRT allows the radiation dose to conform more precisely to the three-dimensional (3-D) shape of the tumor by modulating, or controlling, the intensity of the radiation beam in multiple small volumes. IMRT also allows higher radiation doses to be focused to regions within the tumor while minimizing the dose to surrounding normal critical structures.

2.3.2.1 VMAT

The last type of radiotherapy techniques that is examined is the Volumetric Arc Therapy (VMAT) or RapidArc Radiotherapy Technology. It is an advanced form of IMRT that delivers a precisely sculpted 3D dose

distribution with a 360-degree rotation of the gantry in a single or multi-arc treatment. It is defined as the delivery of a rotational cone beam with variable shape and intensity. VMAT, in less than two minutes, can deliver the dose to the entire tumor in a 360-degree rotation, as opposed to conventional IMRT treatments, during which the machine must rotate several times around the patient or make repeated stops and starts to treat the tumor from several different angles. Thanks to special software and an advanced linear accelerator from Varian VMAT deliver IMRT treatments up to eight times faster than what was in the past; moreover, the algorithm ensures treatment precision, helping to spare surrounding healthy tissue. In a VMAT treatment, the gantry moves continuously, with the MLC leaves and dose rate varying throughout the arc. The TPS computes the dose by sampling the delivery at several discrete gantry angles. In order to create a satisfactory dose plan with a single arc, it is necessary to optimize the field shapes and beam intensities from a large number of gantry angles. However, the field shapes are restricted in that the MLC leaves must be able to move to their new positions within the time required for the gantry to rotate between samples. Unfortunately, the larger the number of sampled gantry angles, the more difficult it is for the TPS to optimize within the MLC leaf motion constraints. A solution was found, employing a technique called progressive sampling to alleviate this problem. Other optimization algorithms have since been developed, and in practice, VMAT is commonly used to refer to any arc therapy technique that uses dose rate variation. VMAT differs from existing techniques like helical IMRT or intensity-modulated arc therapy (IMAT) because it delivers dose to the whole volume, rather than slice by slice. Perhaps the biggest advantage of a VMAT delivery is in its delivery efficiency.

2.4 Plan Quality Assessment

2.4.1 Isodose curve&surface

The Isodose curve is a curve, which connects points of equal dose in a plan, isodose surfaces are surfaces where every point has at least the same dose as the one specified by the user. They are typically used to evaluate the dose distribution in the patient, slice by slice, and the spread of dose out and inside the targets or the OARs. Iso-dose curves and surfaces to investigated the conformity and the absence of relevant hot and cold spots.

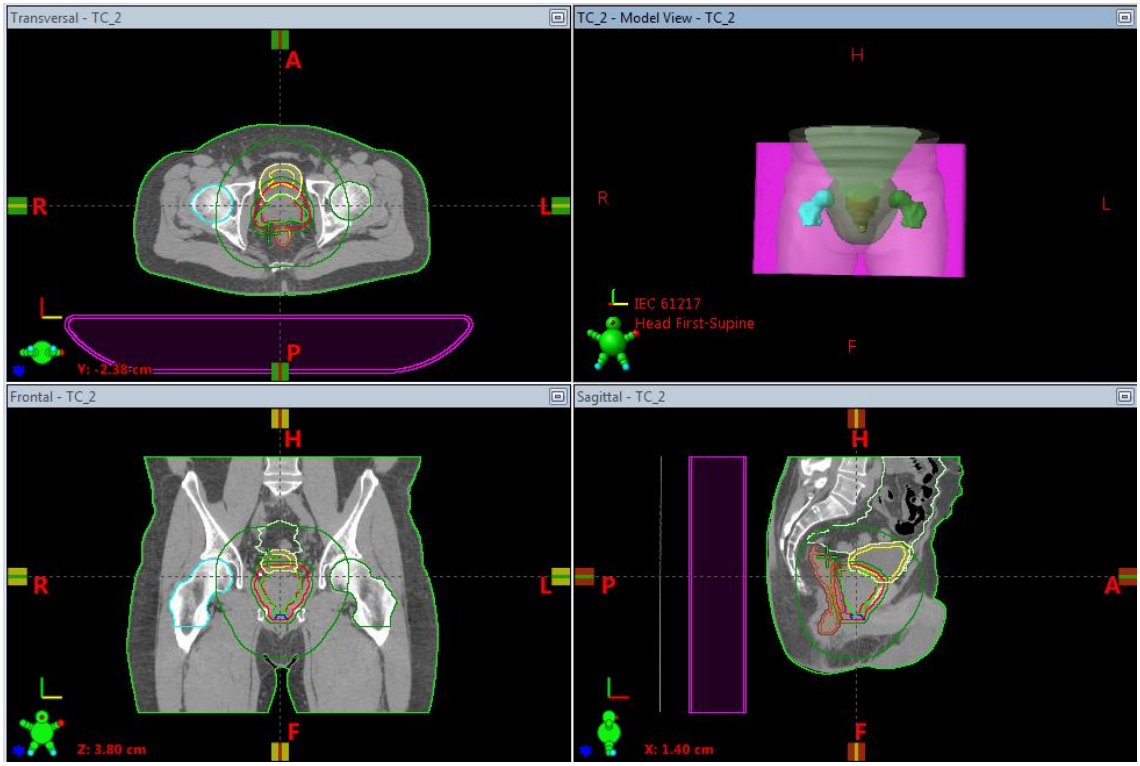


Figure 1: A conformal treatment plan for prostate cancer patients showing isodose curves in the transverse **(A)**, lateral **(B)**, and coronal **(C)** planes. Target volumes and OARs are displayed **(D).**

2.4.2 Dose Volume Histograms (DVH)

To judge the quality of the plans and obtain the plan to be discussed with the radiation oncologists, medical physicists use dose volume histograms (DVHs). Dose volume histograms are histograms where on the x-axis the dose is represented, on the y-axis the volume. They are a useful tool to view the dose coverage in a very compact way. DVHs are automatically calculated in every TPS systems.

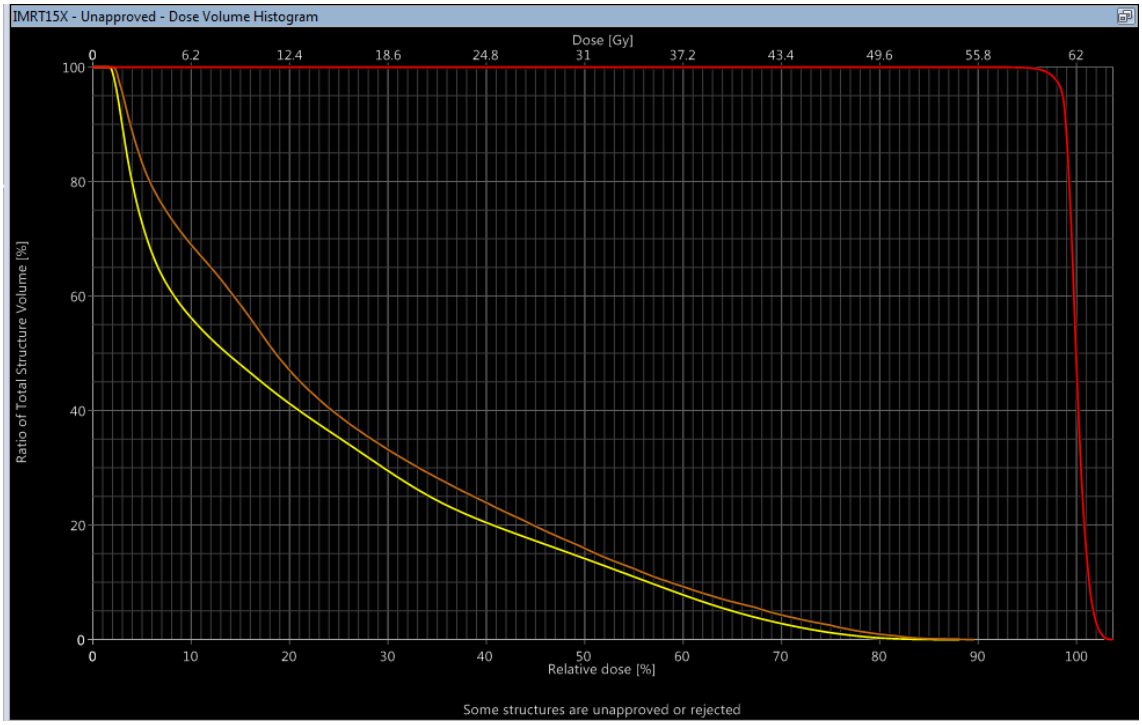


Figure 2: Cumulative dose volume histogram (DVH) is useful in the evaluation process in order to study the level of coverage of each volume at a given dosage.

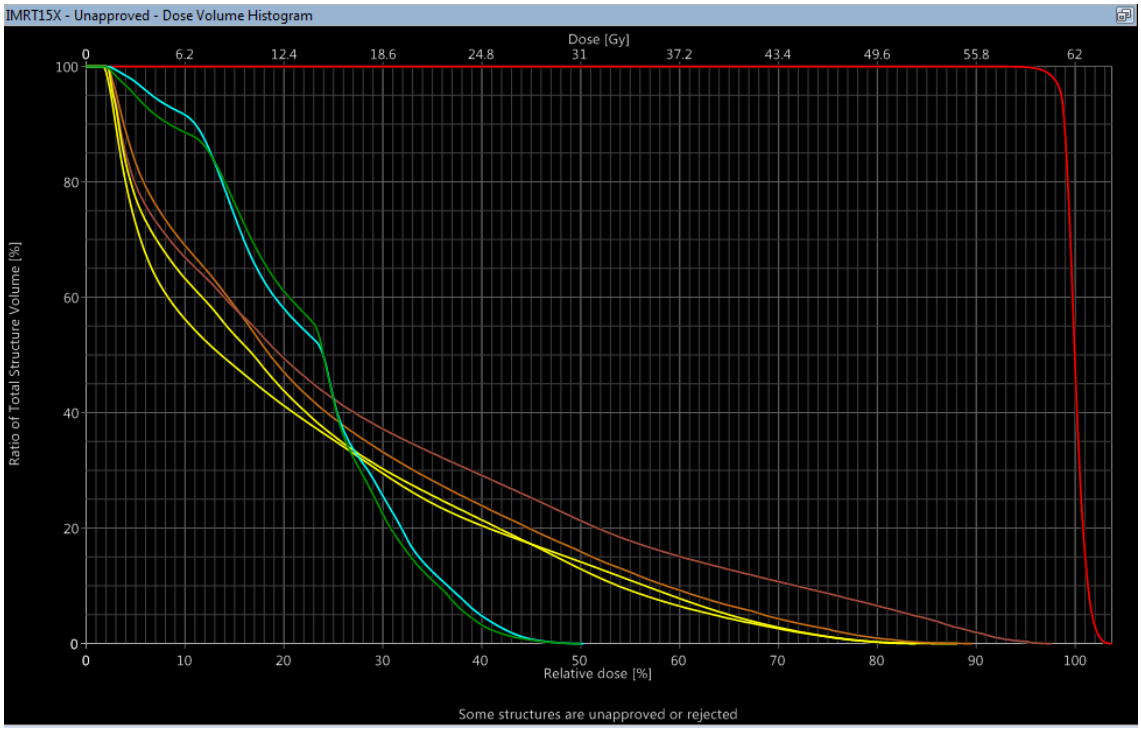


Figure 3: an example of plan comparisons between rival plans is reported

2.4.3 TCP/NTCP

2.5 Radiomics

The term radiomics originates from "radio" which comes from radiology that is the branch of medicine concerning the production and the interpretation of images, and "omics" means that the radiomics has a holistic approach to the study of images encompassing the entire view of the system in study. [d avanzo]

A radiomics study, in order to be conducted requires various steps, described in [lambin] and reported in Figure 4.

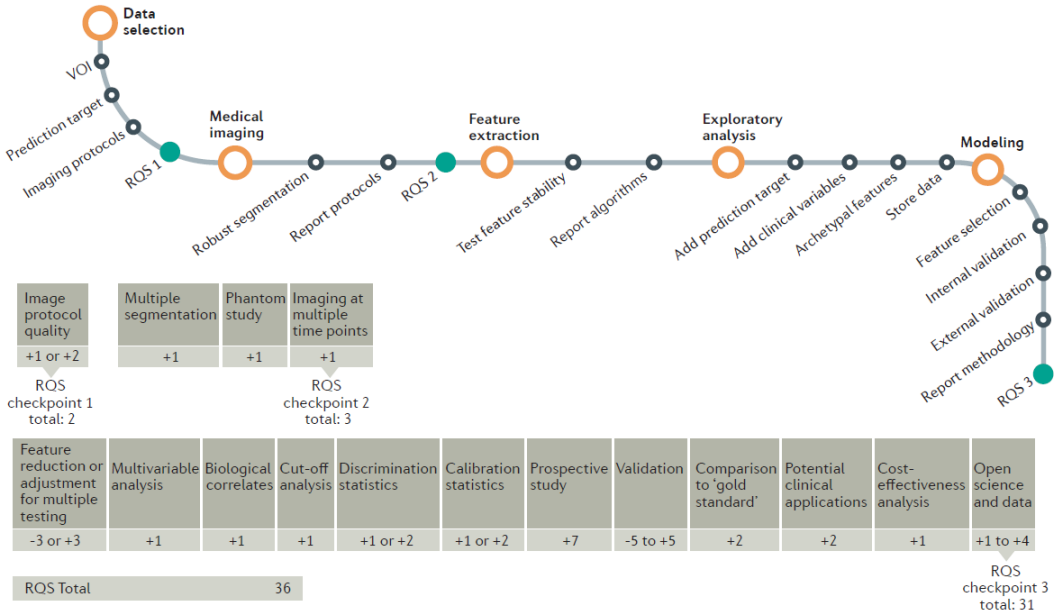


Figure 4: fundamental steps for obtaining robust radiomics features from "**Radiomics: the bridge between medical imaging and personalized medicine", by Lambin et al.**

A brief summary of the typical procedures is given. The first steps concern the implementation of standardized protocols of acquisition for the images, then there is the segmentation, the delineation of the anatomic parts we want to study (this is typically done by a physician). Later the feature extraction part can start. A feature is a descriptor of an image derived from the histogram, texture and shape. After the extraction, there will be a further step, a correlation between the features and biological variables is searched, the main aim is indeed to establish the prognostic power of the features. Further steps are the use of

machine learning techniques to create a mathematical model and relate the value of the features with a biological variable.

Many software are available either open source or commercial, this is mainly due to the fast increasing of the number of papers about radiomics that are published every year has reported in fig. 5.

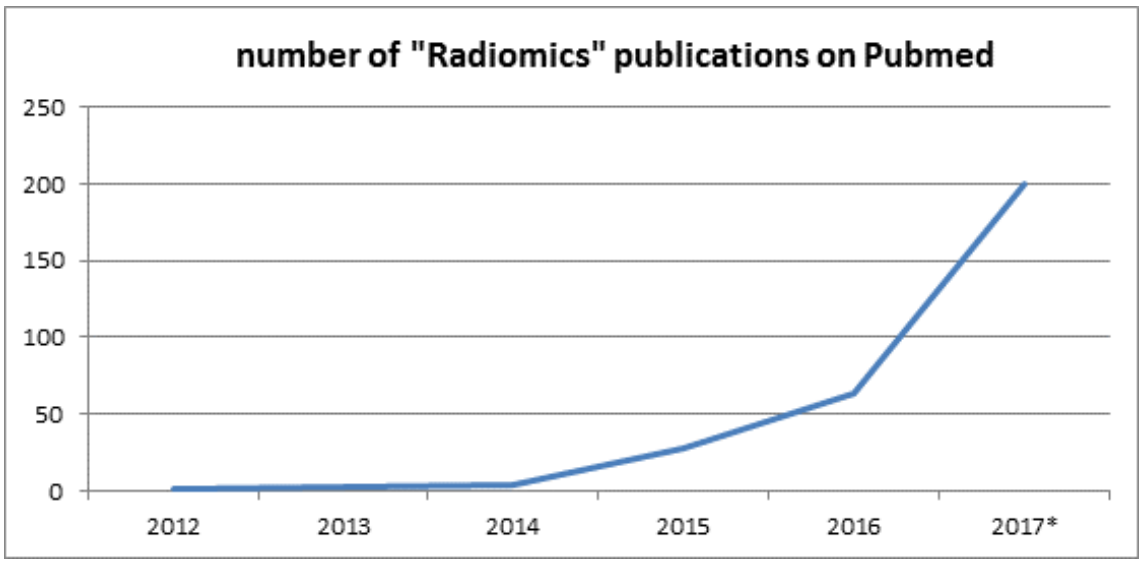


Figure 5: number of papers about radiomics that are published every year

In this work the software, 3DSlicer with the PyRadiomics extension have been used. PyRadiomics is written in python by … and is an open source library, which offers the possibility to extract and the export the data about the extracted features. It allows the analysis of first order and textural features. A summary of these quantities is described in table 1.

$energy$	Energy is a measure of the magnitude of voxel values in an image. A larger value implies a greater sum of the squares of these values.	$\sum_{i=1}^{N_p}(X(i)+c)^2$
$total\ energy$	Total Energy is the value of Energy feature scaled by the volume of the voxel in cubic mm.	$V_{voxel}\sum_{i=1}^{N_p}(X(i)+c)^2$
$entropy$	Entropy specifies the uncertainty/randomness in the image values. It measures the average amount of information required to encode the image values.	$-\sum_{i=1}^{N_g}p(i)\log_2(p(i)+\varepsilon)$
$minimum$	The minimum gray level intensity within the ROI.	$\min(X)$
10th percentile	The 10^{th} percentile of X	
90th percentile	The 90^{th} percentile of X	
$maximum$	The maximum gray level intensity within the ROI.	$\max(X)$
$mean$	The average gray level intensity within the ROI.	$\frac{1}{N_p}\sum_{i=1}^{N_p}X(i)$
median	The median gray level intensity within the ROI.	
$interquartile\ range$	Here $P_{25}\ e\ P_{75}$ are the 25^{th} and 75^{th} percentile of the	$P_{75}-P_{25}$

	image array, respectively.	
$range$	The range of values	$\max(X) - \min(X)$
MAD	Mean Absolute Deviation (MAD) is the mean distance of all intensity values from the Mean Value of the image array.	$\frac{1}{N_p}\sum_{i=1}^{N_p}\lvert X(i) - \bar{X}\rvert$
$rMAD$	Robust Mean Absolute Deviation (rMAD) is the mean distance of all intensity values from the Mean Value calculated on the subset of image array with gray levels in between, or equal to the 10^{th} and 90^{th} percentile. Root Mean Squared	$\frac{1}{N_{10-90}}\sum_{i=1}^{N_{10-90}}\lvert X_{10-90}(i) - \bar{X}_{10-90}\rvert$
RMS	(RMS) is the square root of the mean of all the squared intensity values. It is another measure of the magnitude of the image values.	$\sqrt{\frac{1}{N_p}\sum_{i=1}^{N_p}(X(i)+c)^2}$
$standard\ deviation$	Standard Deviation measures the amount of variation or dispersion from the Mean Value.	$\sqrt{\frac{1}{N_p}\sum_{i=1}^{N_p}(X(i)-\bar{X})^2}$
$skewness$	Skewness measures the asymmetry of the distribution of values about the Mean value.	$\frac{\mu_3}{\sigma^3} = \frac{\frac{1}{N_p}\sum_{i=1}^{N_p}(X(i)-\bar{X})^3}{\left(\sqrt{\frac{1}{N_p}\sum_{i=1}^{N_p}(X(i)-\bar{X})^2}\right)^3}$

$kurtosis$	Kurtosis is a measure of the 'peakedness' of the distribution of values in the image ROI.	$\frac{\mu_4}{\sigma^4} = \frac{\frac{1}{N_p}\sum_{i=1}^{N_p}(X(i)-\bar{X})^4}{\left(\frac{1}{N_p}\sqrt{\sum_{i=1}^{N_p}(X(i)-\bar{X})^2}\right)^2}$
$variance$	Variance is the mean of the squared distances of each intensity value from the Mean value.	$\frac{1}{N_p}\sum_{i=1}^{N_p}(X(i)-\bar{X})^2$
$uniformity$	Uniformity is a measure of the sum of the squares of each intensity value. This is a measure of the heterogeneity of the image array.	$\sum_{i=1}^{N_p} p(i)^2$

Table 1.summary of first order radiomics features

2.5.1 First Order Features

Let:

- X be a set of N_p voxels included in the ROI
- P(*i*) be the first order histogram with N_g discrete intensity levels, where N_g is the number of non-zero bins, equally spaced from 0 with a width defined in the binWidth parameter.
- *p(i)* be the normalized first order histogram and equal to P(*i*) N_p

2.6 Data analysis

For the statistical analysis, the package R has been used.

2.6.1 Analysis of the distribution of treatment plans on selected cases

A Principal Component Analysis (PCA) is introduced for the analysis of the distribution of treatment plans. It is about an algorithm that reduces the dimensionality of data while retaining most of the variation in the dataset. It identifies linearly independent combinations of parameters that summarize the statistical correlations present in the data. PCA is an orthogonal linear transformation, or else, a new orthogonal coordinate system is defined by the principal components. You have that the greatest variance within the data occurs along the first coordinate (the first principal component); the second-greatest variance occurs along the second coordinate, and so on. The first component, that represents the most important source of variation in the data, is defined as the eigenvector with the largest eigenvalue. Otherwise, the last component represents the least important process contributing to the variation. Next to PCA, it has defined the Exact Logistic Regression; the correlations between PC coefficients and toxicity are assessed using exact logistic regression. Exact logistic regression is used to model binary outcome variables in which the log odd of the outcome is modeled as a linear combination of the predictor variables. It is used when the sample size is too small for a regular logistic regression, which uses the standard maximum-likelihood-based estimator.

Chapter 3

Results

The investigated plans for patients are shown in figure 6-8.

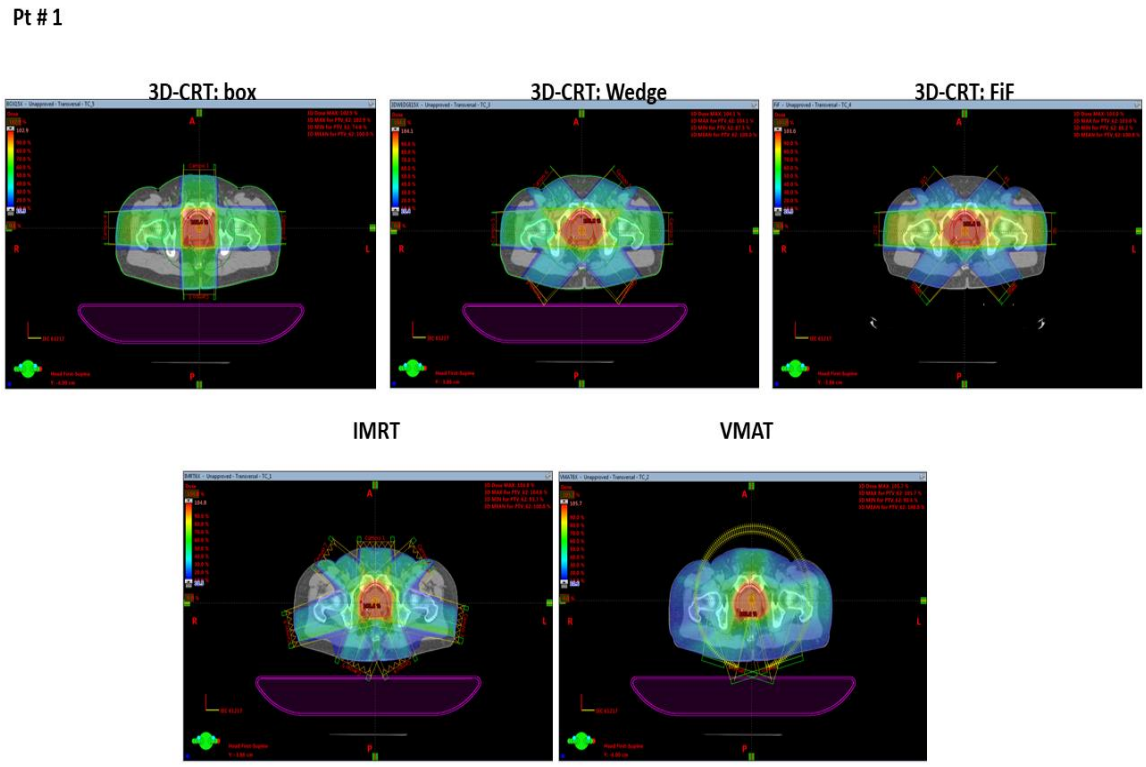


Figure 6: the investigated plans for patient # 1 are shown with the indication of the type of plan

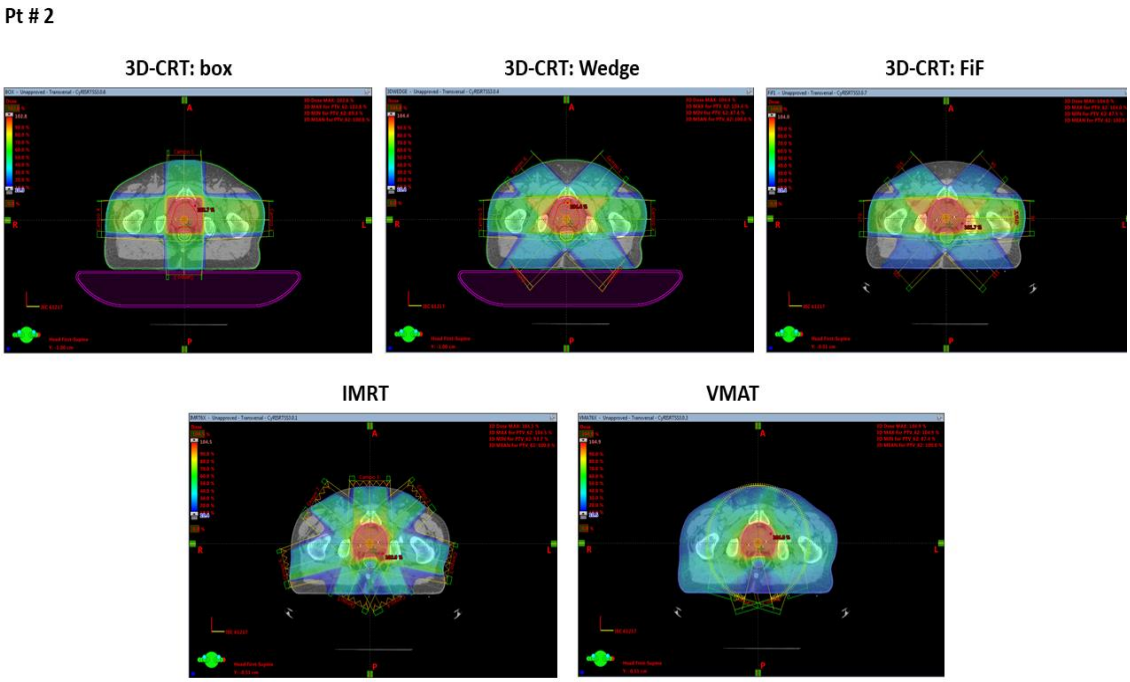


Figure 7: the investigated plans for patient # 2 are shown with the indication of the type of plan

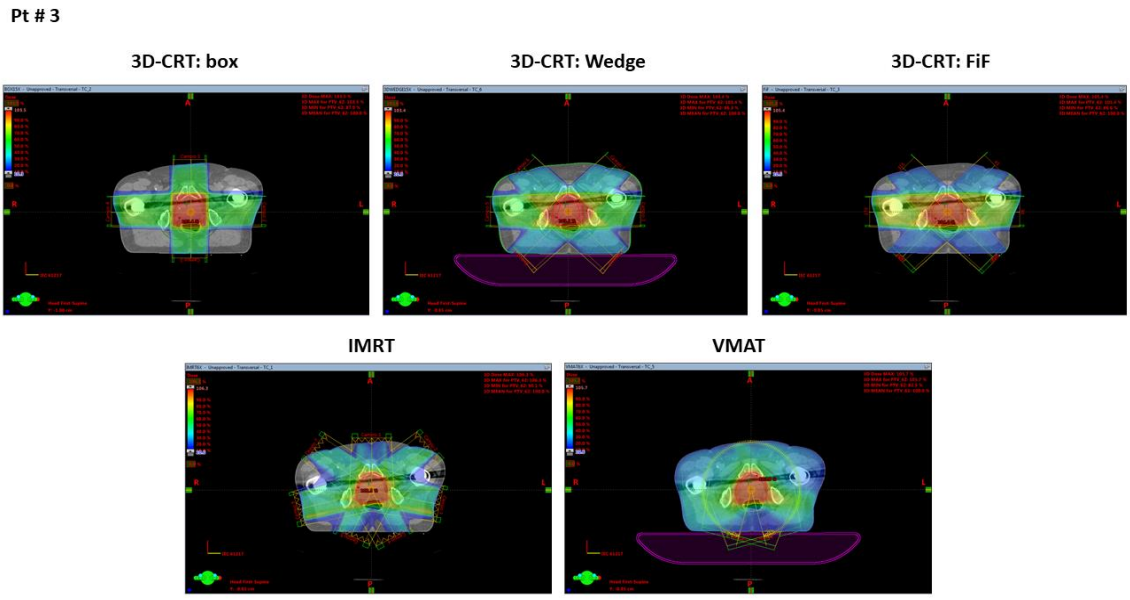


Figure 8: the investigated plans for patient # 3 are shown with the indication of the type of plan

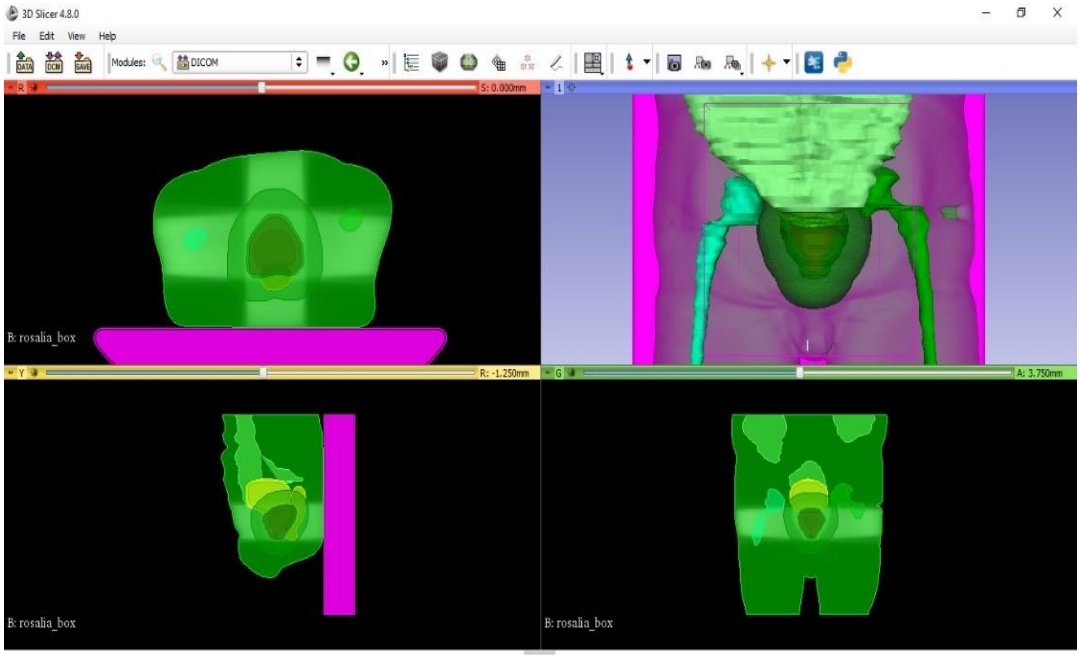


Figure 9: the imported dose of box plan for one of patients in slicer

When the dose is imported in slicer (e.g. figure 9) the dose distribution can be appreciated see figure 9 and first order features can be calculated.

3.1 Analysis of the distribution of treatment plans on selected cases

TCP

TCP values were calculated for PTV62 (only for comparative aim) from the planned and accumulated differential DVHs i.e. $\{D_i.V_i\}$ according to the Poisson formalism [16.17]:

$$TCP = exp\left\{-\sum_{i=1}^{M} V_i exp\left[\ln(N_0) - \alpha D_i\left(1 + \frac{D_i/n_f}{\alpha/\beta}\right)\right]\right\} \quad \text{(eq.2)}$$

where V_i is the fraction volume receiving the dose D_i (with $\sum_i^M V_i = 1$), α and β are the radiobiological parameters (with α =0.25/Gy and α/β=10Gy, see table 2) and n_f is the number of fractions.

NTCP

For bladder and rectum, NTCP values were calculated by using the following:

$$NTCP = \frac{1}{\sqrt{2\pi}} \int_{--}^{t} \exp(-\frac{t^2}{2}) dt \quad \text{(eq. 3)}$$

where t is defined as reported in [20]

$$t = \frac{D - TD_{50}(v)}{mTD_{50}(v)} \quad \text{(eq.4)}$$

Here D is the maximum dose expressed in terms of normalized total dose at 2 Gy / fraction (NTD2), m is the slope of the NTCP curve versus the dose.

In the equation 4, v is the effective dose from DVH. To take into consideration the heterogeneous dose distribution (i.e. the whole DVHs), v was calculated as follows [20]:

$$v = \sum_{i=1}^{M} V_i \left[\frac{D_i \left(D_i / n_f + a / b \right)}{D_{\max} \left(D_{\max} / n_f + a / b \right)} \right] \quad \text{(eq.5)}$$

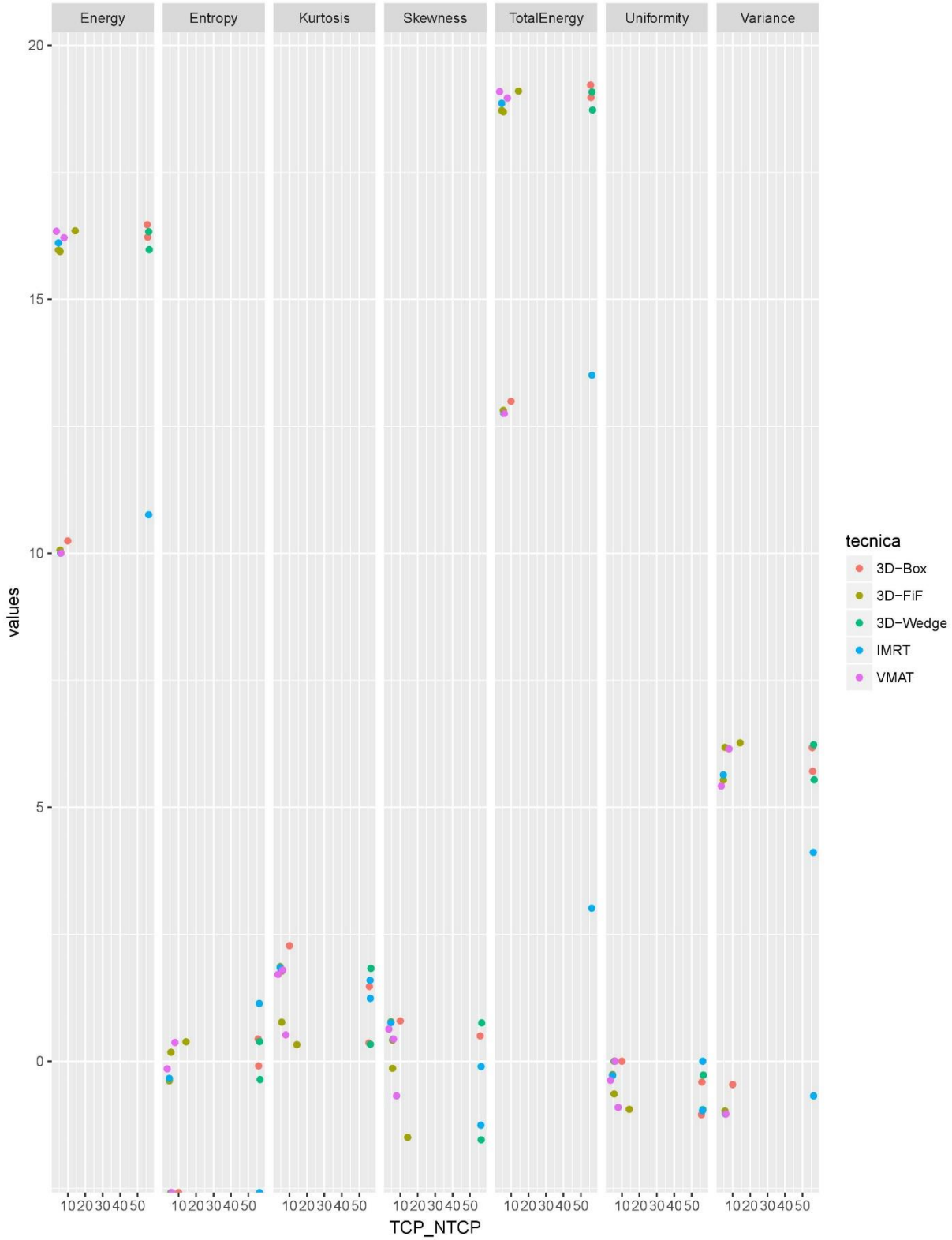
and $D_{\max}$ is the maximal dose to the analyzed organ.

$TD_{50}(v)$ is defined as follows:

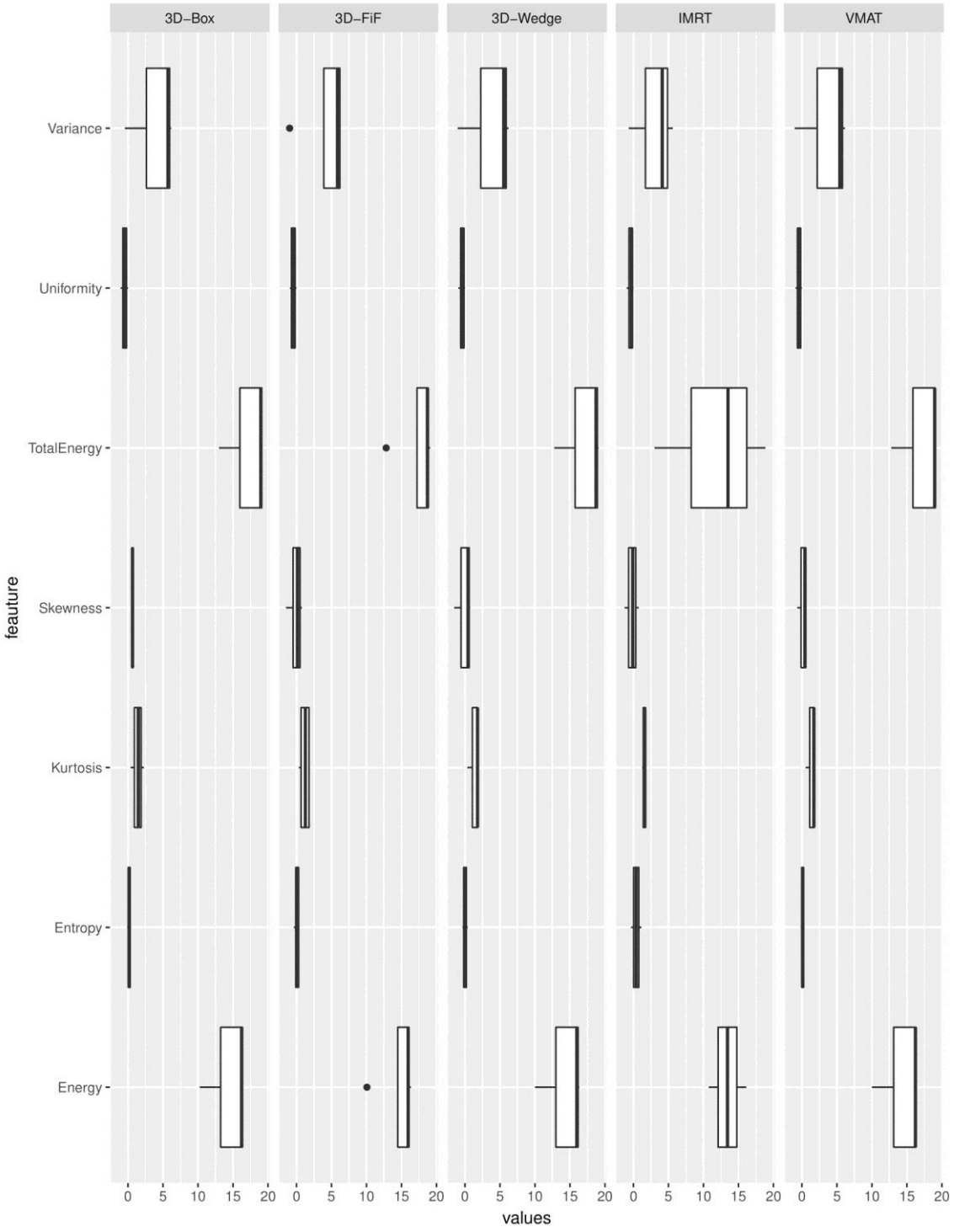
$$TD_{50}(v) = TD_{50}(1) \cdot v^{-n} \quad \text{(eq.6)}$$

where $TD_{50}(1)$ is the tolerance dose to the whole organ, leading to a 50% complication probability and n is a parameter related to the organ response to radiotherapy. Used parameters for bladder and rectum are reported in table XX.

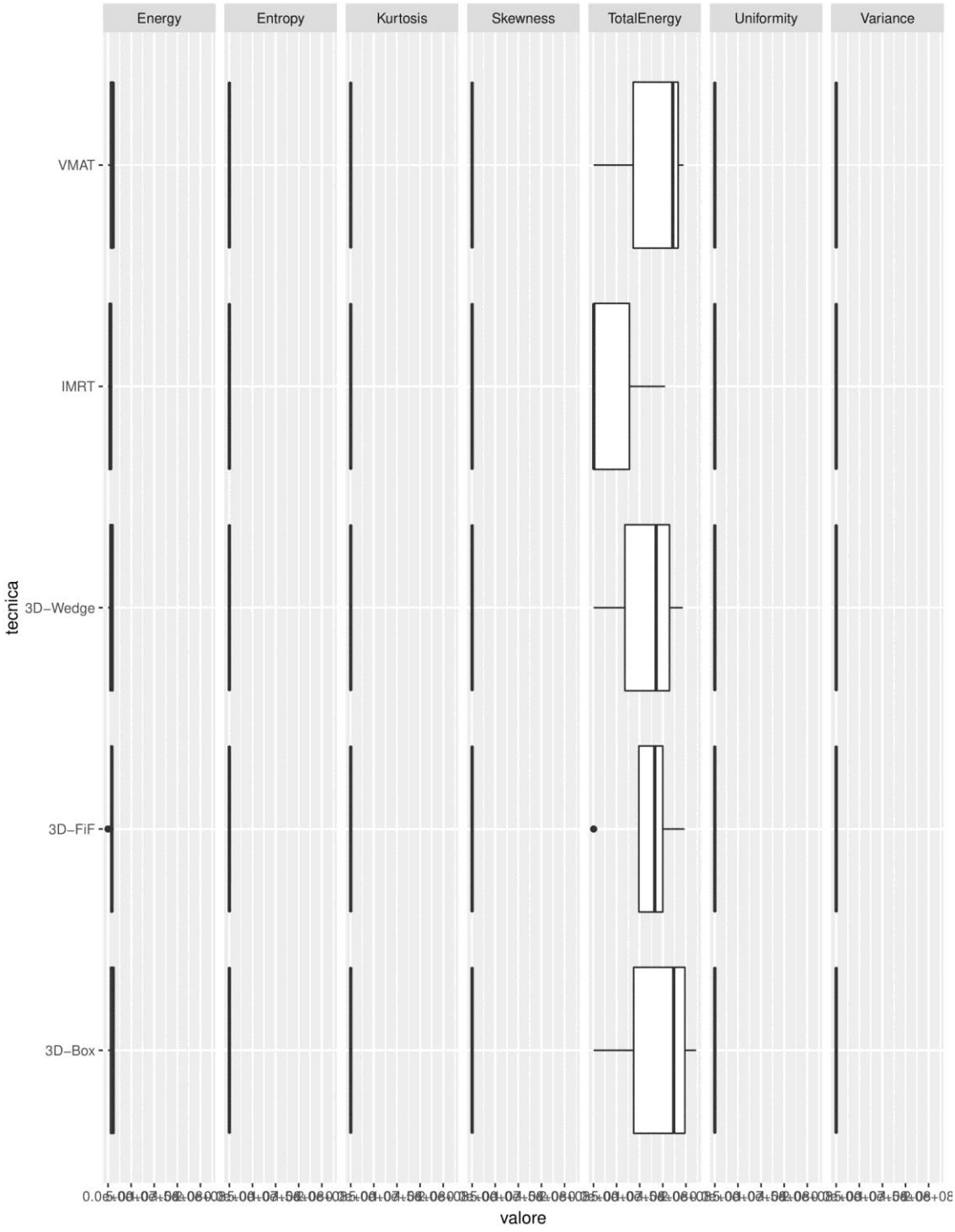
BLADDER_NTCP



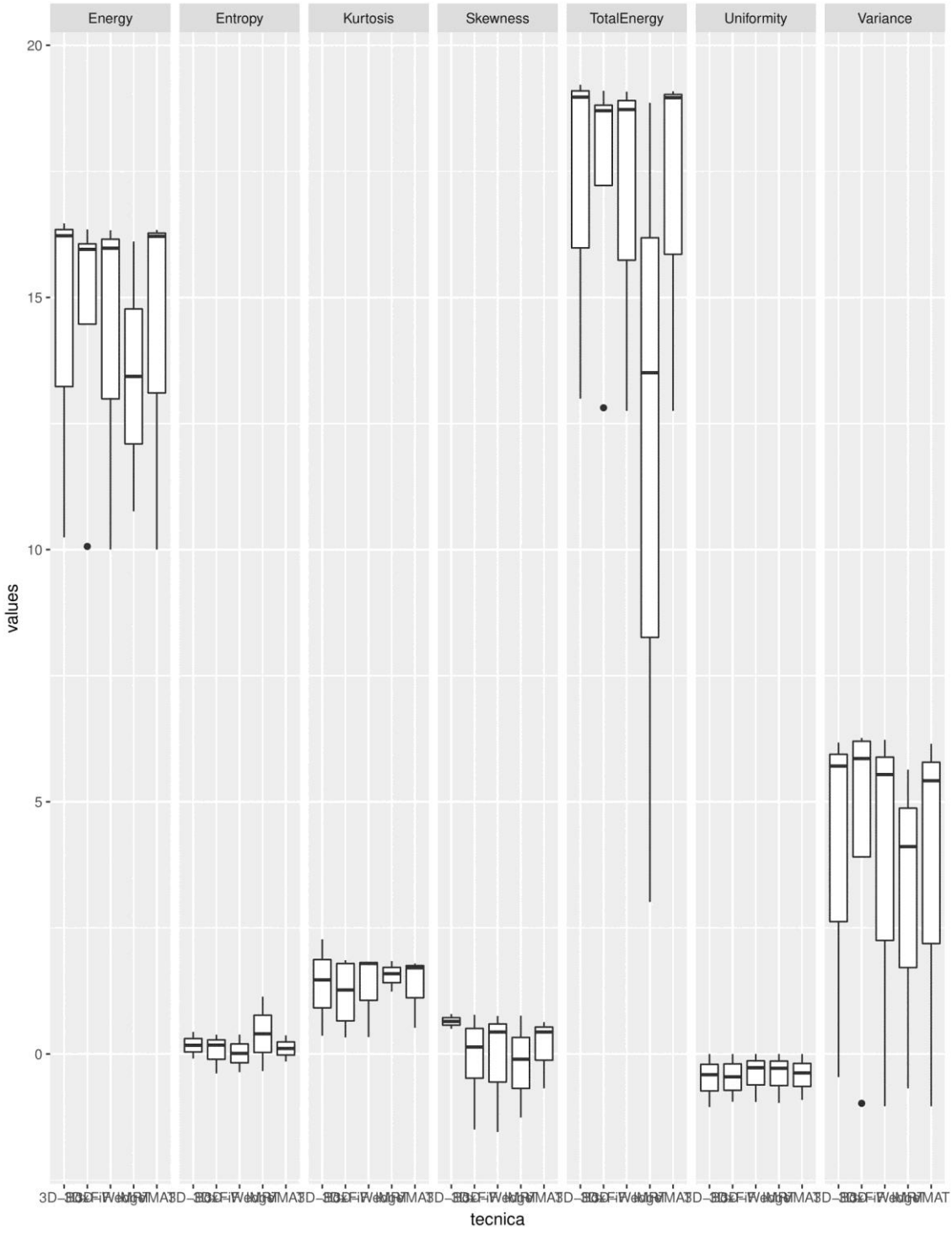
BLADDER_FEATURE



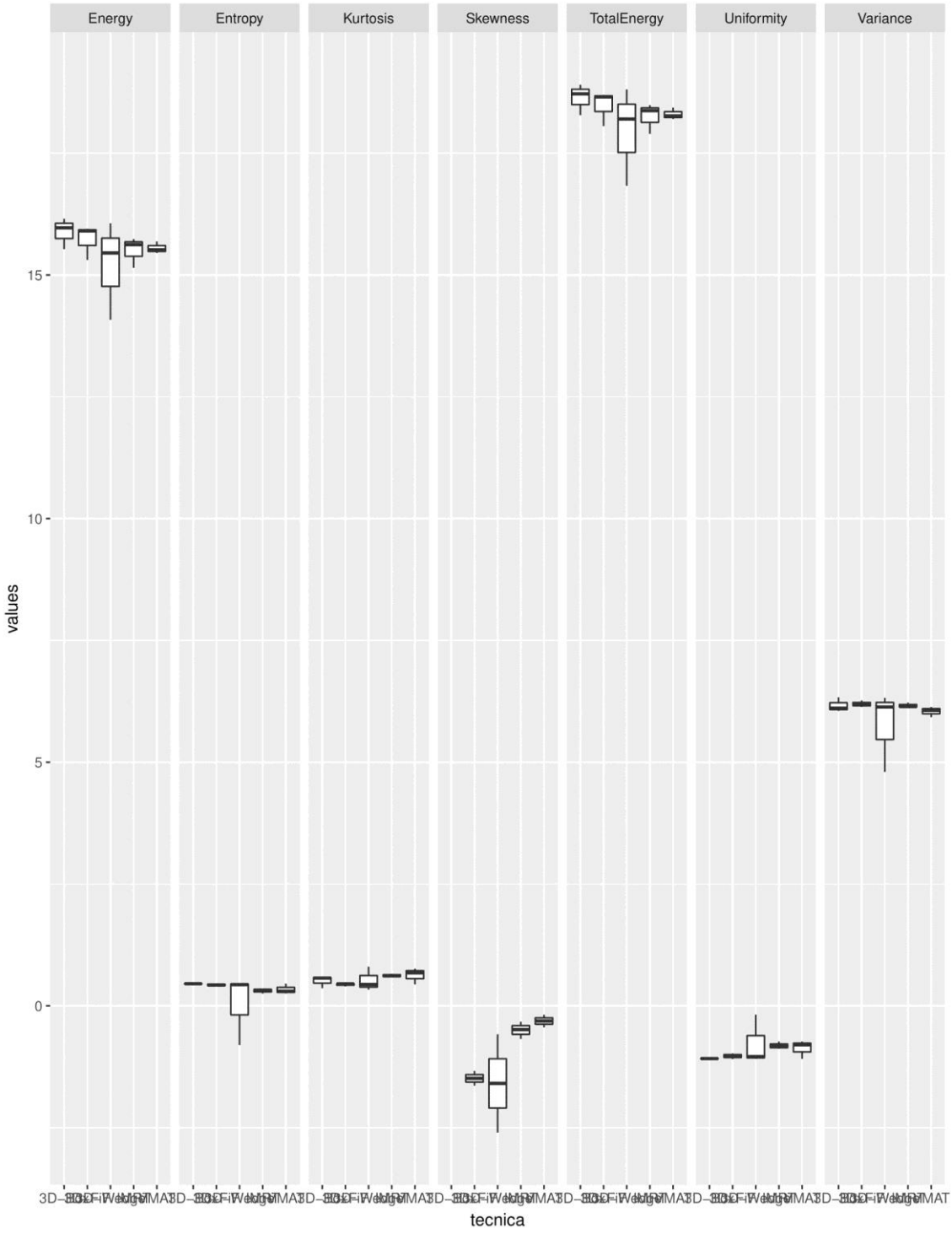
BLADDER_VALORE



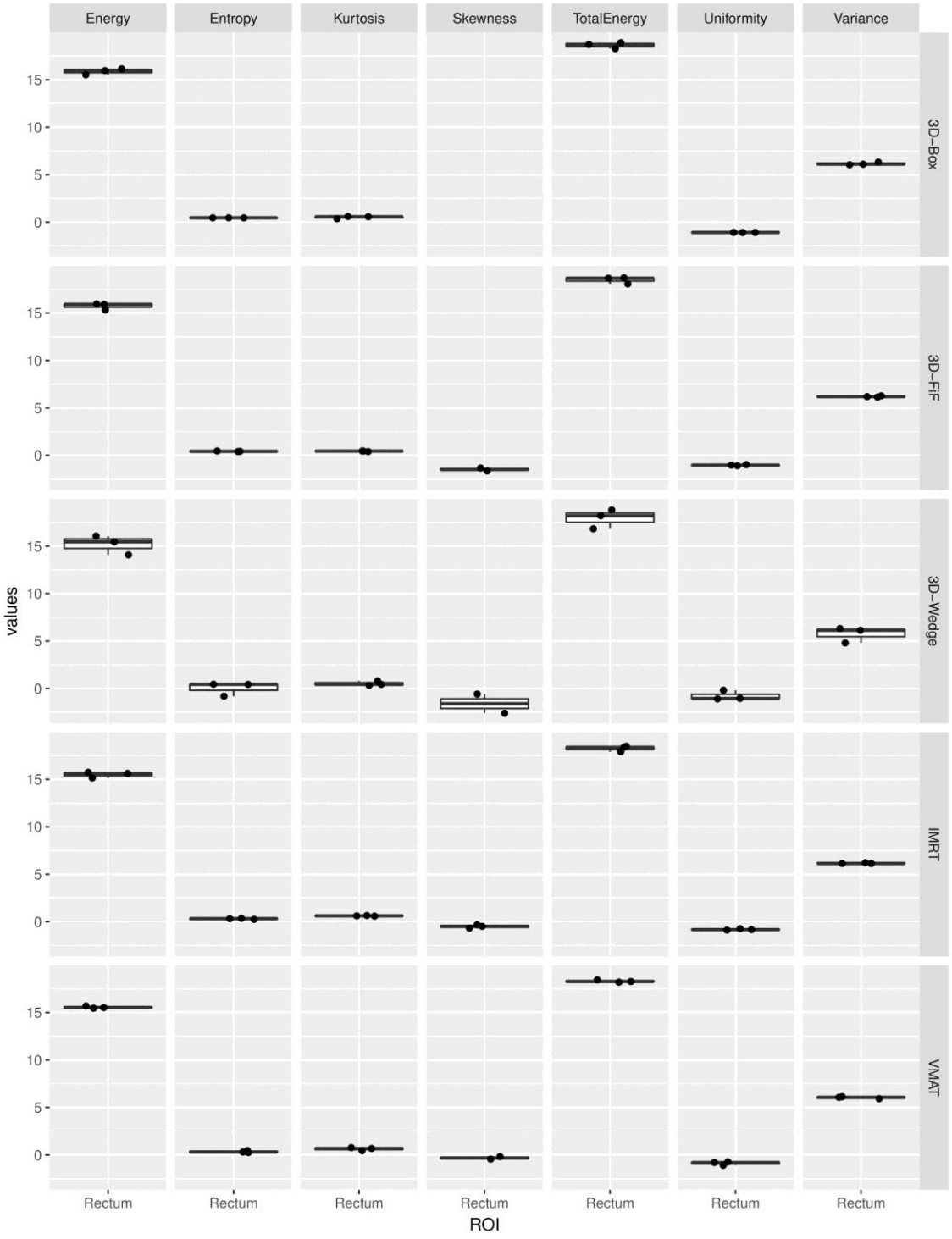
BLADDER_VALUES



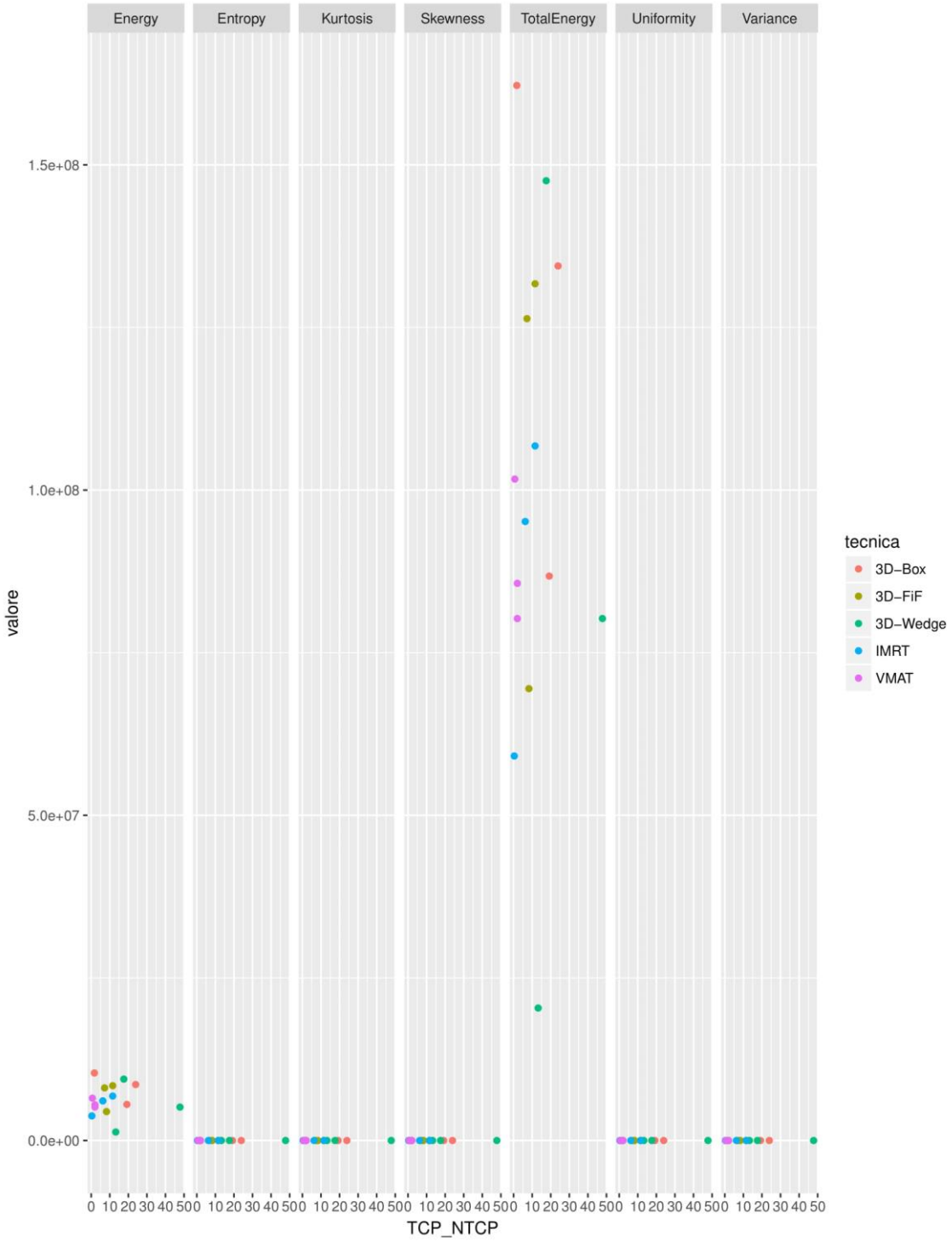
RECTUM_FEATURE



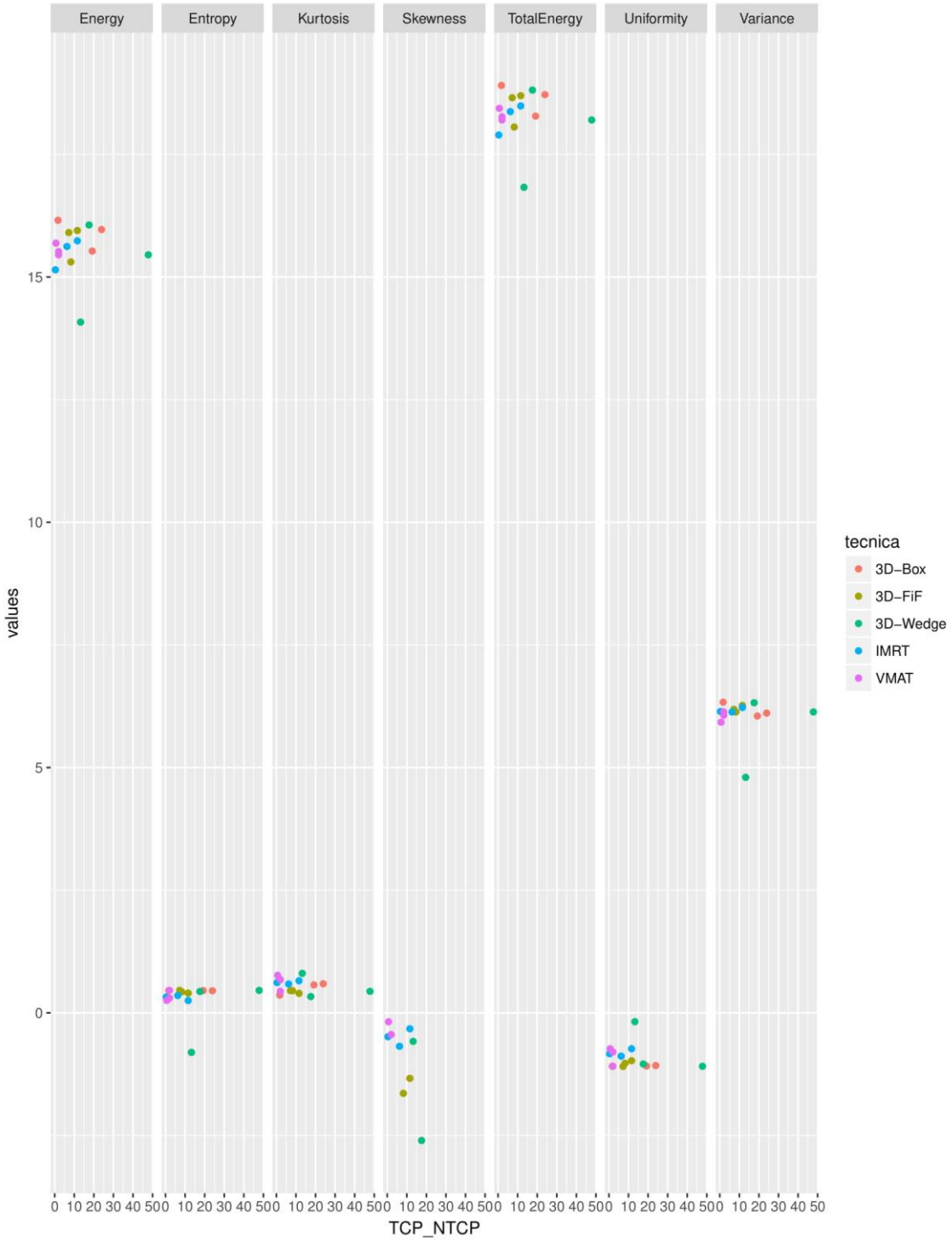
RECTUM_ROI


Energy
Entropy
Kurtosis
Skewness
TotalEnergy
Uniformity
Variance
3D−Box
3D−FiF
3D−Wedge
IMRT
VMAT
values
15
10
5
0
Rectum
ROI

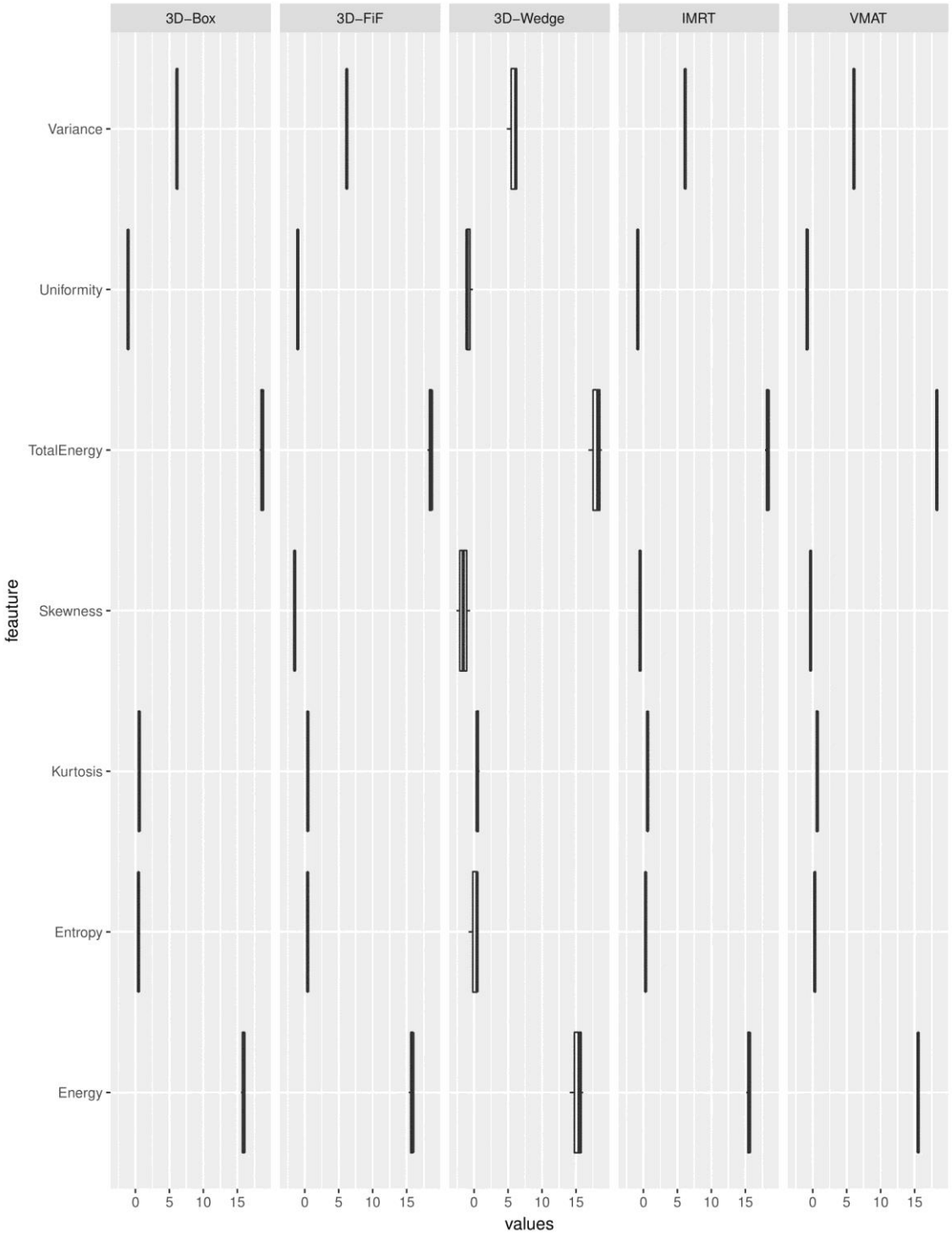
RECTUM_TCP_NTCP



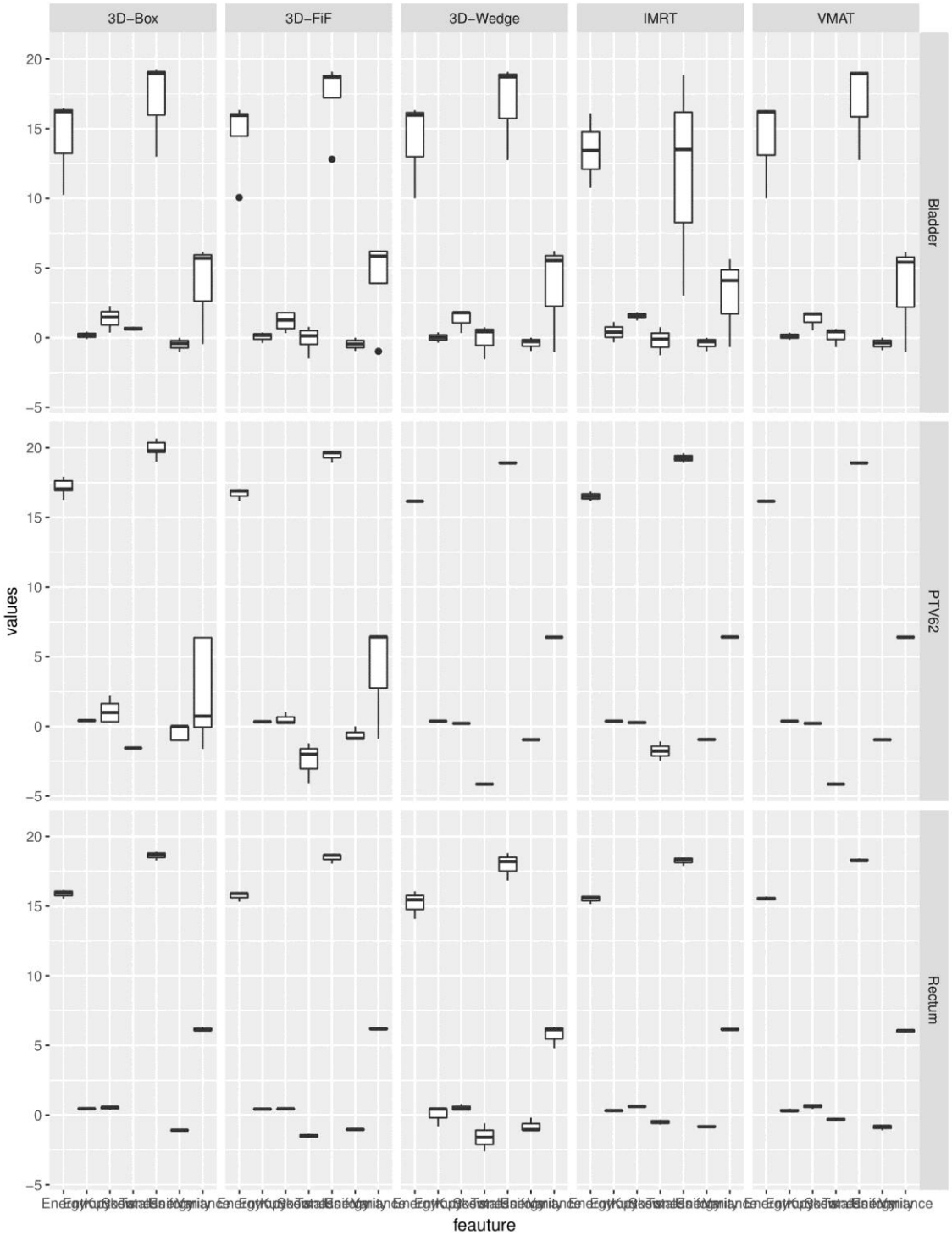
RECTUM_TCP_NTCP



RECTUM_VALUES



TECNICA_ROI


3D−Box
3D−FiF
3D−Wedge
IMRT
VMAT
Bladder
PTV62
Rectum
values
feauture

3.2 Comparison of rival treatment plans (developed for the same patient / district)

The evaluation of the Auto-planning tool has been performed in 3 cases in order to validate this approach in a clinical setting.

NCTP and TCP values for each radiotherapy technique have been computed and compared with those obtained by manual planning.

Moreover, specific radiomic features have been extracted for each treatment technique using a dedicated software to better characterize the illness.

Chapter 4

Conclusion

Thanks to the use of different radiotherapy techniques we investigated the dose distribution applied for prostate cancer. We used radiomics to identify a more appropriate indicator to understand the more appropriate technique for this type of cancer. Different techniques have been tested. Radiation treatment planning is designed to maximize benefits and keep potential risks to a minimum. This involves working out an exact site, an angle of radiation, and an optimal dose.

Bibliography

[1] Deasy J 1996 Poisson formulas for tumor control probability with clonogenic proliferation Radiat. Res. **145**, 382–4

[2] L Strigari, M D'Andrea, A Abate, M Benassi "A heterogeneous dose distribution in simultaneous integrated boost: the role of the clonogenic cell density on the tumor control probability" Phys Med Biol **53**, 5257–73 (2008)

[3] L Strigari, G Arcangeli, S Arcangeli, M Benassi. "Mathematical model for evaluating incidence of acute rectal Toxicity during conventional or hypofractionated radiotherapy courses for prostate cancer" Int J Rad Onc Biol Phys **73**, 1454-60 (2008)