Physics and Imaging in Radiation Oncology 14 (2020) 87–94

Contents lists available at ScienceDirect

Physics and Imaging in Radiation Oncology

journal homepage: www.elsevier.com/locate/phro

Original Research Article

Associations between voxel-level accumulated dose and rectal toxicity in prostate radiotherapy

Leila E.A. Shelley[a,b,d,*], Michael P.F. Sutcliffe[a,d], Simon J. Thomas[a,c], David J. Noble[a,e], Marina Romanchikova[a,h], Karl Harrison[a,g], Amy M. Bates[a,f], Neil G. Burnet[a,i], Raj Jena[a,e]

[a] Cancer Research UK VoxTox Research Group, Cambridge University Hospitals NHS Foundation Trust, Department of Oncology, Addenbrooke's Hospital, Cambridge CB2 0QQ, United Kingdom
[b] Edinburgh Cancer Centre, Western General Hospital, Edinburgh EH4 2XU, United Kingdom
[c] Department of Medical Physics and Clinical Engineering, Cambridge University Hospitals NHS Foundation Trust, Addenbrooke's Hospital, Cambridge CB2 0QQ, United Kingdom
[d] Department of Engineering, University of Cambridge, Trumpington St, Cambridge CB21PZ, United Kingdom
[e] Department of Oncology, University of Cambridge, Cambridge Biomedical Campus, Addenbrooke's Hospital, Hills Road, Cambridge CB2 0QQ, United Kingdom
[f] Cambridge Clinical Trials Unit, Cambridge University Hospitals NHS Foundation Trust, Addenbrooke's Hospital, Cambridge CB2 0QQ, United Kingdom
[g] Cavendish Laboratory, University of Cambridge, J J Thomson Avenue, Cambridge CB3 0HE, United Kingdom
[h] National Physical Laboratory, Teddington TW11 0JE, United Kingdom
[i] University of Manchester, Manchester Academic Health Science Centre, Manchester M13 9PL, United Kingdom

ARTICLE INFO

Keywords:
Rectal toxicity
Finite element modelling
Dose-surface maps
Delivered dose
Prostate cancer
Adaptive radiotherapy

ABSTRACT

Background and Purpose: Associations between dose and rectal toxicity in prostate radiotherapy are generally poorly understood. Evaluating spatial dose distributions to the rectal wall (RW) may lead to improvements in dose-toxicity modelling by incorporating geometric information, masked by dose-volume histograms. Furthermore, predictive power may be strengthened by incorporating the effects of interfraction motion into delivered dose calculations.

Here we interrogate 3D dose distributions for patients with and without toxicity to identify rectal subregions at risk (SRR), and compare the discriminatory ability of planned and delivered dose.

Material and Methods: Daily delivered dose to the rectum was calculated using image guidance scans, and accumulated at the voxel level using biomechanical finite element modelling. SRRs were statistically determined for rectal bleeding, proctitis, faecal incontinence and stool frequency from a training set (n = 139), and tested on a validation set (n = 47).

Results: SRR patterns differed per endpoint. Analysing dose to SRRs improved discriminative ability with respect to the full RW for three of four endpoints. Training set AUC and OR analysis produced stronger toxicity associations from accumulated dose than planned dose. For rectal bleeding in particular, accumulated dose to the SRR (AUC 0.76) improved upon dose-toxicity associations derived from planned dose to the RW (AUC 0.63). However, validation results could not be considered significant.

Conclusions: Voxel-level analysis of dose to the RW revealed SRRs associated with rectal toxicity, suggesting non-homogeneous intra-organ radiosensitivity. Incorporating spatial features of accumulated delivered dose improved dose-toxicity associations. This may be an important tool for adaptive radiotherapy in the future.

1. Introduction

Rectal toxicity remains a clinical issue in prostate radiotherapy. Existing toxicity prediction models may no longer be appropriate for the modern high-dose gradient complex treatments currently being delivered, as the historic data used to derive constraints for dose-limiting organs were based on 3D conformal techniques [1,2]. The underlying mechanisms and pathophysiology of rectal toxicity remain relatively poorly understood [3,4].

Dose-volume histogram (DVH) constraints are commonly applied in radiotherapy treatment planning to achieve an optimal compromise between tumour coverage and healthy organ sparing. However, it is widely acknowledged that DVH parameters are limited by their lack of spatial dose information [5–7]. Studies into more appropriate dose

* Corresponding author at: Edinburgh Cancer Centre, Western General Hospital, Edinburgh EH4 2XU, United Kingdom.
E-mail address: leila.shelley@nhslothian.scot.nhs.uk (L.E.A. Shelley).

https://doi.org/10.1016/j.phro.2020.05.006
Received 11 March 2020; Received in revised form 15 May 2020; Accepted 18 May 2020

descriptors for the rectal wall (RW) have generally been focused on parameterisation of dose surface maps (DSMs) where the RW is virtually cut and unfolded to a normalised 2D map so that geometric dose information can be preserved [5,7–16].

An advancing area of research, which complements interrogation of localised dose features, is voxelwise dose-toxicity analysis [17,18]. Investigating associations between voxel-level dose and toxicity can improve the accuracy in identifying heterogeneous areas of heightened dose sensitivity, or rectal subregions at risk (SRRs) [18]. One solution offering voxel-resolution anatomical modelling and 3D tracking of organ motion, is biomechanical finite element (FE) analysis [19–24]. FE modelling applies constitutive equations to anatomical simulations, and unlike the DSM method, is not restricted to normalised in-plane expansion. Additionally, by comparison with image intensity-based deformable registration tools which can struggle at large deformations and are generally also limited to planar expansions [25], FE offers superior spatial functionality.

An overarching limitation and recommendation of studies examining dose-toxicity associations in prostate radiotherapy, has been that motion-inclusive dosimetric data are required for a more complete understanding of the underlying mechanisms [4]. Radiation treatments to the pelvic region are particularly susceptible to interfraction motion which can lead to large deviations between planned dose and that actually delivered. This introduces an inherent uncertainty in dose-toxicity modelling based on planned dose.

Previous work [16] presented the calculation and accumulation of delivered dose to DSMs of the RW. The 2D DSMs were resampled such that the number of pixels spanning the width was equal to the rectal length from the planning scan [7,8]. However, when applied to dose accumulation, the DSM approach was limited in that modelling of rectal expansion and deformation was in-plane only and uniformly normalised about the circumference of the rectum. The motivation for implementing biomechanical FE modelling into the VoxTox workflow was to provide a more anatomically representative basis for voxel-level accumulation of delivered dose [26]. FE models have previously been developed for the rectum in prostate radiotherapy [22,23], but to date, dose accumulation studies have focused on the liver [21].

The aim of this study was to determine rectal SRRs from planned and accumulated FE-DSMs, and investigate associations with toxicity.

2. Material and methods

2.1. VoxTox study design & patient information

The VoxTox research programme is an observational study linking radiation dose to toxicity outcomes [27]. It received approval from the National Research Ethics Service (NRES) Committee East of England (13/EE/0008) in February 2013 and is part of the UK Clinical Research Network Study Portfolio (UK CRN ID 13716). Appropriate consent was obtained from all participants.

Data were analysed for a set of 186 prostate cancer patients recruited to the VoxTox study. This 'consolidation' cohort consisted of patients prescribed 74 Gy in 37 fractions (n = 110, 59%) and 60 Gy in 20 fractions (n = 76, 41%), treated with TomoTherapy® (Accuray, Sunnyvale, CA), and whose baseline characteristics and follow-up data were formally and prospectively recorded. Planning dose constraints were applied following the CHHiP trial protocol [28,29]. Doses for patients prescribed 60 Gy in 20 fractions were radiobiologically converted to the equivalent dose in 37 fractions using an alpha/beta ratio of 2.1 Gy based on an analysis of rectal toxicity in the CHHiP trial [28,29]. Equivalent levels of cumulative toxicity incidence were assumed between the two prescription arms.

All patients were treated with image-guided intensity modulated radiotherapy (IG-IMRT), with a daily megavoltage computed tomography (MVCT) scan acquired immediately prior to treatment delivery for online target localisation [30]. According to local imaging protocols, if rectal dilation was deemed excessive, remedial action was taken prior to delivering the radiation treatment. Kilovoltage (kV) planning scans were manually contoured, and the rectum was defined from the rectosigmoid junction to the most inferior slice containing both ischial tuberosities [31]. For daily MVCT scans, the rectum was identified using a locally developed autosegmentation system based on the Chan-Vese algorithm [32]. Manual contouring of the 4174 MVCT scans (over 62,000 slices) in this study would not have been practicable, and the autosegmentation system has previously been shown to fall within an acceptable range for intra- and inter-observer variability [33]. Dose was calculated directly from the MVCT images (and recalculated for kVCT images) using CheckTomo, an independent ray-tracing dose calculation algorithm [34,35]. Where the MVCT field of view was shorter than the kVCT, missing data were substituted using planned dose [16].

2.2. Toxicity data

Toxicity data were prospectively collected using study-specific electronic clinical reporting forms. These were designed for robust collection of toxicity data, and the raw data were used to populate recognised systems. Mapping rules were externally validated. Here, we focused on toxicity endpoints affecting the patient's quality of life, commonly reported in the literature [6,7,10,11,13,15,18,36], and present results for rectal bleeding ⩾G2, proctitis ⩾G2, faecal incontinence ⩾G1, all Common Terminology Criteria for Adverse Events (CTCAE) v4.03 [37] and stool frequency ⩾G1, Late Effects on Normal Tissue: Subjective, Objective, Management (LENT/SOM) scale [38]. Patients were included where a minimum of two years' follow-up data were available, with complete late toxicity follow-up history recorded at 6, 12, and 24 months. Cumulative incidence at 2 years was investigated, and event rates are shown in Table 1.

2.3. Training and validation sets

Patients were split into training (n = 139, 75%), and validation (n = 47, 25%) sets. In order to retain the same ratios of toxicity incidence as the whole cohort, the composition of patients within training and validation sets was uniquely generated for each individual toxicity endpoint. The data were non-randomly split by stratifying first by fractionation, then by toxicity incidence. Patients were then randomised into either training or validation set. This maintained toxicity rates between each group as much as possible, as shown in Table 1.

Table 1
Two-year toxicity incidence rates for full cohort, training (75%) and validation (25%) sets, stratified by prescription regime. RB = rectal bleeding, proc = proctitis, incont = faecal incontinence, freq = stool frequency.

Toxicity endpoint	Total cohort n (%)	74 Gy n (%)	60 Gy n (%)	Training n (%)	74 Gy n (%)	60 Gy n (%)	Validation n (%)	74 Gy n (%)	60 Gy n (%)
	186 (100)	110 (59)	76 (41)	139 (75)	82 (59)	57 (41)	47 (25)	28 (60)	19 (40)
RB	21 (11)	10 (5)	11 (6)	16 (12)	8 (6)	8 (6)	5 (11)	2 (4)	3 (6)
Proc.	27 (15)	13 (7)	14 (8)	20 (14)	9 (6)	11 (8)	7 (15)	4 (9)	3 (6)
Incont.	32 (17)	21 (11)	11 (6)	24 (17)	16 (12)	8 (6)	8 (17)	5 (11)	3 (6)
Freq.	45 (24)	28 (15)	17 (9)	34 (24)	21 (15)	13 (9)	11 (23)	7 (15)	4 (9)

2.4. Finite element modelling

A physical-based rectal simulation was created in the Abaqus Computer Aided Engineering (CAE) FE Analysis environment (Dassault Systèmes) using 30 elements around the rectal circumference and 80 along the rectal length. This was approximately equivalent to the width dimensions used in the previous DSM analysis [16], but with increased resolution along the length of the rectum in order to improve the accuracy of tracking extra-planar motion. The model was able to deform and expand in 3D according to material properties, loading and boundary conditions. The boundary conditions were defined by the rectal contours, manually defined on kVCT or autosegmented on MVCT images. Displacement loading was applied such that the model expanded based on the shape and position of the contours, rather than pressure loading where defining parameters can be difficult due to challenges in experimental verification [26]. Further details are provided in Supplementary Material 1. When each simulation was complete, the dose was determined in MATLAB (MathWorks®, Natick, MA) at each voxel by interpolation of the 3D dose matrix calculated in CheckTomo. The use of a common starting point for each simulation allowed voxels to be tracked, and dose histories accumulated. This facilitated like-for-like comparisons between patients, as well as between planned and delivered dose. The resulting dose surface maps generated using FE modelling will be termed FE-DSMs.

2.5. Identifying rectal subregions at risk

Rectal SRRs were determined using the training set for each toxicity endpoint following a similar approach to that of Dréan et al. [18]. For each voxel of the FE-DSM, a two-sample t-test for equality of means was performed in MATLAB. The two groups were defined by patients with and without toxicity and the null hypothesis was that there was no difference in mean voxel dose between groups. Voxels were identified in areas where the dose was significantly higher for patients with toxicity than those without. Thresholds were determined based on an unadjusted p-value < 0.05, as well as p-values adjusted for false discovery rates using the Benjamini-Hochberg procedure [39]. Where the p-value at each voxel was less than the threshold, the dose difference between patients with and without toxicity was considered significant, and the null hypothesis was rejected. From the resulting voxelwise p-value map, a binary mask was generated whereby a value of 1 was assigned when the null hypothesis was rejected, and a value of 0 was assigned otherwise. Post-processing was performed on the binary mask to remove clusters of fewer than 10 pixels (unless clusters would have been connected had it not been for the posterior splitting of the FE-DSM for visualisation in 2D), to smooth sharp edges and discontinuities using a least-squares smoothing filter (Savitzky-Golay filter in the MATLAB image processing toolbox), and to fill in any remaining holes. This affected fewer than 5% of pixels for all endpoints other than rectal bleeding which was up to 10% due to the proportionately larger SRR and detected holes. The final SRRs for both unadjusted and adjusted thresholds are overlaid on the p-value maps for each endpoint shown in Fig. 1 (for binary masks, see Supplementary Material 2).

The SRR binary masks were applied to each patient's planned and accumulated FE-DSM such that only the dose to voxels from the non-zero regions were considered. The equivalent uniform dose (EUD) [40] was then calculated (using $a = 11.1$ [2]) based on the dose to voxels within these subregions. EUD of the full RW was also calculated, as this was previously found to be associated with rectal bleeding ⩾grade 1 and proctitis using 2D DSMs[16]. Receiver operator characteristic (ROC) analysis was performed in SPSS (IBM, 23.0.0.2) and the area under the curve (AUC) was reported with 95% confidence intervals (CIs). Odds ratios (ORs) were also calculated using logistic regression. The analysis was repeated by applying the SRR binary masks to the validation set (which had been kept blind to the process of generating SRRs). Validation AUC and ORs were calculated in order to compare with training set results.

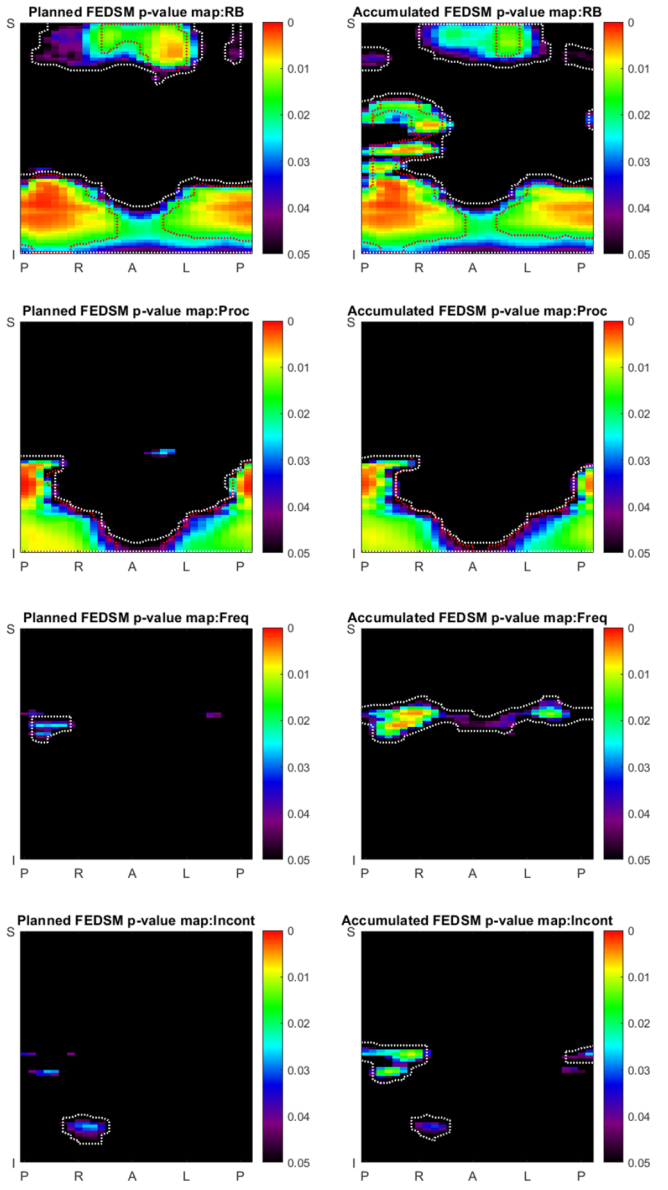


Fig. 1. Map of p-values for rectal bleeding (RB), proctitis (proc), stool frequency (freq), and faecal incontinence (incont), indicating regions of significant dose difference between patients with and without toxicity for planned and accumulated dose to the rectal wall. White dashed line represents SRRs where $p < 0.05$, red dashed line indicates adjusted p-value SRR, where applicable. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3. Results

3.1. Localised dose response in subregions at risk

The p-value maps shown in Fig. 1 indicate the SRRs of the RW where significant dose difference were observed between patients with and without toxicity. Spatial patterns of localised dose response differed per endpoint. The relative proportions of SRR area (from unadjusted p-value maps) with respect to the full RW were: rectal bleeding 47/56%, proctitis 22/22%, faecal incontinence 3/6%, stool frequency 2/11%, from planned/accumulated dose. For rectal bleeding, the SRRs spanned the circumference of the inferior RW and most of the superior RW for both planned and accumulated FE-DSMs, with additional expansion into the mid-right-posterior RW on the accumulated FE-DSM. The SRRs associated with rectal bleeding were much larger in area than those

Table 2
Equivalent uniform dose (EUD) comparison between full rectal wall (RW) and subregions at risk (SRR) for planned and accumulated dose, split by toxicity groups for the training set. Standard deviations (SD) and EUD differences (Δ_{tox}) are presented between patients with and without toxicity alongside corresponding *p*-values. Differences between planned and accumulated dose for RW and SRR had $p<0.001$ for all endpoints. RB = rectal bleeding, proc = proctitis, incont = faecal incontinence, freq = stool frequency.

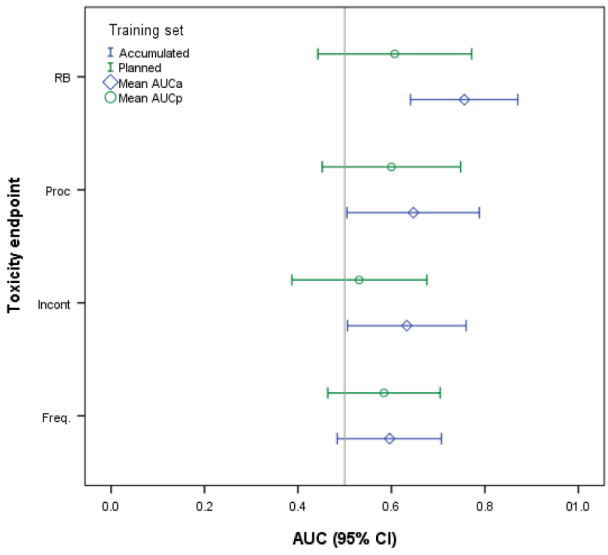
Endpoint	Dose	EUD_{RW}	$SD(EUD_{RW})$	$\Delta_{tox}(EUD_{RW})$	*p*-value	EUD_{SRR}	$SD(EUD_{SRR})$	$\Delta_{tox}(EUD_{SRR})$	*p*-value
RB	Planned	60.9	2.5	1.2	0.006	51.4	7.2	2.7	0.2
	Accumulated	59.8	3.0	1.7	0.003	54.0	4.7	3.8	0.002
Proc	Planned	61.0	2.3	0.8	0.2	39.5	9.2	3.8	0.1
	Accumulated	60.0	2.7	0.9	0.2	43.1	7.2	3.2	0.06
Incont	Planned	61.0	2.4	0.2	0.8	31.9	17.4	3.0	0.4
	Accumulated	60.0	2.9	0.2	0.8	51.4	6.0	2.7	0.05
Freq	Planned	60.9	2.4	0.4	0.4	42.0	8.1	2.3	0.2
	Accumulated	60.0	2.9	0.7	0.2	58.7	5.7	1.7	0.1

found for other toxicities. For proctitis, SRRs were observed from the mid-to-inferior posterior RW and connected by a band spanning the circumference inferiorly, common to both planned and accumulated FE-DSMs. For faecal incontinence, a cluster at the inferior right-lateral RW was common to both planned and accumulated FE-DSMs, with an additional mid-posterior to mid-right cluster on the accumulated FE-DSM. For stool frequency, a single cluster was observed in the right-posterior mid-RW for planned dose, and for accumulated dose, a larger cluster spanning the full circumference was located in the mid-upper RW. By comparison, adjusted p-value maps resulted in reduced, more focussed SRR regions for rectal bleeding and proctitis, and did not produce a result for faecal incontinence or stool frequency. Results discussed below refer to unadjusted p-value SRRs unless otherwise specified.

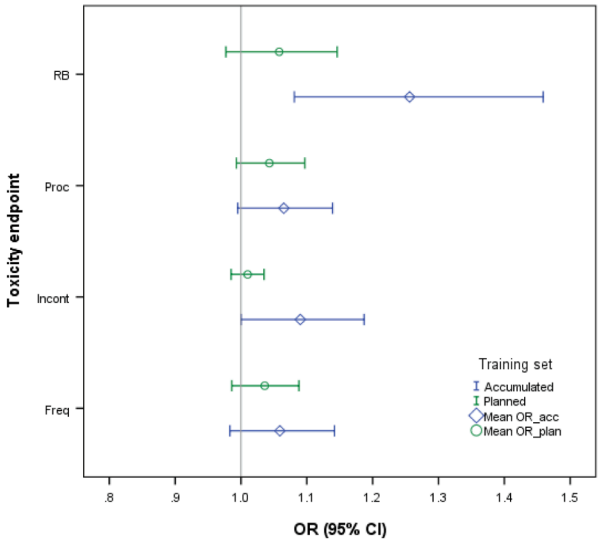
3.2. Dose-toxicity associations

When comparing EUD of the full RW, accumulated dose was lower for all four endpoints by an average of 1 Gy (range −0.9, −1.1; $p < 0.001$) with respect to planned dose (Table 2). The difference in EUD between patients with and without toxicity was greater for accumulated dose for three of four endpoints, but differences could only be considered significant for rectal bleeding. In contrast, when focusing on SRR regions, EUD was greater for accumulated dose than planned dose by an average of 10.6 Gy (range 2.6, 19.5; $p < 0.001$). The difference in EUD between patients with and without toxicity was greater for SRR than RW. When splitting SRR EUD by toxicity group, the difference was greater from planned dose for three of four endpoints, but the standard deviation was reduced for accumulated dose compared with planned dose by an average of 4.6 Gy (range −2.0, −11.4). However, the only EUD differences considered significant between toxicity groups were for accumulated dose to SRRs for rectal bleeding and faecal incontinence.

AUCs and ORs for SRR dose-toxicity analysis (with corresponding 95% CIs) are shown in Fig. 2. Results for full RW and adjusted p-values are included alongside a summary of these results in Table 3 (95% CIs included in Supplementary material 3). When comparing AUCs for the full RW, only rectal bleeding was found to have AUC > 0.6, and accumulated dose was more strongly associated than planned dose. For planned dose, there was no advantage to using EUD to SRRs rather than the full RW in terms of discriminative ability. For rectal bleeding, proctitis, and faecal incontinence, AUCs from accumulated SRRs were greater than those from the RW by 7%, 4%, and 12%, respectively. Results for stool frequency could not be considered significant. For SRRs, AUCs were greater when investigating associations from accumulated dose with respect to planned dose (rectal bleeding +15%, proctitis +5%, faecal incontinence +10%). For these endpoints, accumulated dose results had AUC > 0.6 with min 95% CI > 0.5,



(a)



(b)

Fig. 2. (a) Area under the receiver operator characteristic curve (AUC), and (b) odds ratios (OR), with corresponding 95% confidence intervals (CI), for the training set (n = 139). RB = rectal bleeding, proc = proctitis, incont = faecal incontinence, freq = stool frequency.

Table 3
Dose-toxicity analysis for planned and accumulated equivalent uniform dose (EUD) to the full rectal wall (RW) and rectal subregion at risk (SRR). Area under the ROC curve (AUC) and odds ratios (OR) were calculated to compare training (*t*) and validation (*v*) sets. Results for SRRs based on p-values adjusted for false discovery rates were also included (AUC_{FDR}). Confidence intervals (CI) for SRR results are shown in Fig. 2 and 3, and are included for all results in Supplementary material 3. RB = rectal bleeding, proc = proctitis, incont = faecal incontinence, freq = stool frequency.

Endpoint	FE-DSM	AUC_{RW}(t)	AUC_{SRR}(t)	AUC_{FDR}(t)	AUC_{SRR}(v)	OR_{SRR}(t)	OR_{FDR}(t)	OR_{SRR}(v)
RB	Planned	0.63	0.61	0.61	0.58	1.06	1.05	1.04
	Accumulated	0.69	0.76	0.80	0.63	1.26	1.28	1.12
Proc	Planned	0.61	0.60	0.58	0.45	1.04	1.14	0.95
	Accumulated	0.61	0.65	0.63	0.39	1.07	1.06	0.93
Incont	Planned	0.54	0.53	–	0.45	1.01	–	0.99
	Accumulated	0.51	0.63	–	0.49	1.09	–	1.02
Freq	Planned	0.58	0.58	–	0.49	1.04	–	0.97
	Accumulated	0.60	0.60	–	0.58	1.06	–	0.98

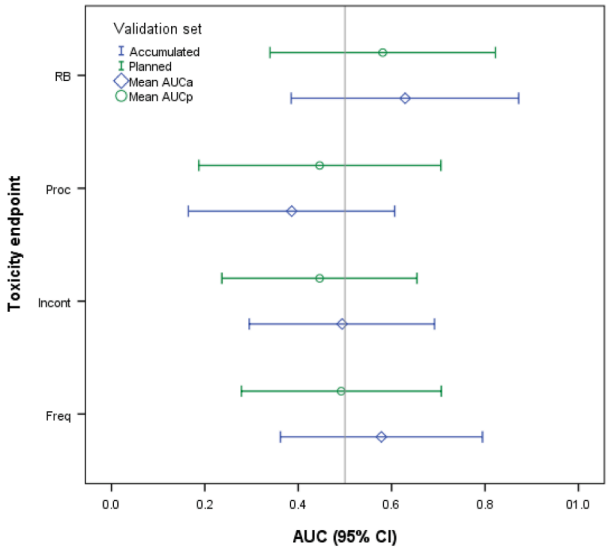
indicating stronger associations than corresponding planned dose results (where min 95% CI < 0.5). OR results indicated that accumulated dose was more strongly associated with rectal bleeding and incontinence than planned dose, by +20% and +8%, respectively. However, the lower bounds of the 95% CIs were < 1 for proctitis, stool frequency, and all planned dose results. For the two endpoints yielding SRRs from p-value maps after adjustment for false discovery rates, rectal bleeding AUC for accumulated dose improved by 4% and for planned dose was unchanged, whereas for proctitis the AUC reduced by 2–3% for both planned and accumulated dose. Differences were not significant compared to unadjusted results.

Validation results, as shown in Fig. 3, were inherently lower with larger CIs than corresponding training results as expected due to the smaller sample size. Due to the widths of the CIs, the validation results could not be considered sufficiently significant to verify the findings from the training set.

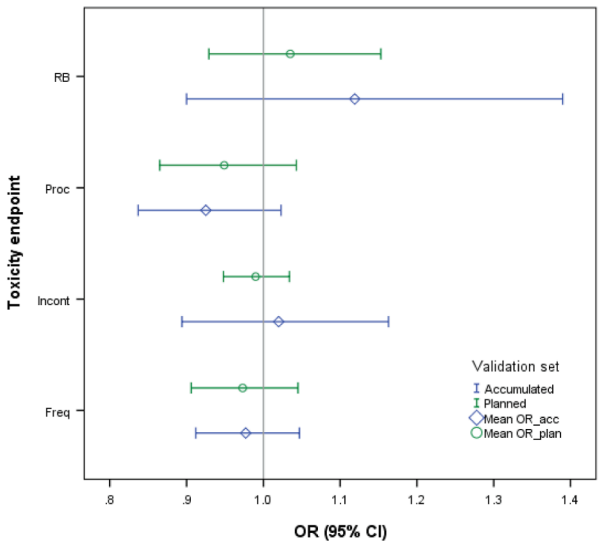
4. Discussion

The investigation of associations between localised dose to rectal SRRs and toxicity is an area of increasing interest. In the current study, subregions of the RW were identified by statistical evaluation of voxel-level dose differences between FE-DSMs of patients with and without toxicity. The addition of biomechanical FE modelling provided a more anatomically representative solution than the previous 2D-based normalised DSM approach. SRRs were generated based on planned and accumulated FE-DSMs, indicating spatial clusters and localised dose patterns unique to each toxicity endpoint. These findings develop upon previous work investigating spatial aspects of rectal DSMs with regards to improving toxicity prediction [16,27]. Several studies have used DSM analysis to demonstrate links between toxicity and mid-to-intermediate dose levels (i.e. other than maximum or prescription level doses) [5,7,10–13,15]. This suggests that further knowledge and understanding of the interaction between dose level, spatial distribution and location within the RW may improve endpoint-specific toxicity prediction.

From the SRR analysis presented here, of particular note is that although patterns differed per endpoint, regions of interest were generally identified away from the highest dose levels (i.e. closest to the prostate), and extended towards lower dose levels, posteriorly, superiorly and/or inferiorly, depending on the particular endpoint. This effect has been observed in previous studies, with several authors reporting associations between dose to the inferior rectum and rectal bleeding [8,10,12,13,18]. Dréan et al. [18] conducted voxel-level DVH analysis of planned dose to the rectal volume, and found rectal bleeding ⩾grade 1 to be associated with dose to a SRR in the inferior-anterior hemi-rectum. Onjukka et al. [10] evaluated planned dose to the



(a)



(b)

Fig. 3. (a) Area under the receiver operator characteristic curve (AUC). and (b) odds ratios (OR), with corresponding 95% confidence intervals (CI), for the validation set (n = 47). RB = rectal bleeding, proc = proctitis, incont = faecal incontinence, freq = stool frequency.

combined RW and anal canal using DSMs, and found inferior regions corresponding to the anal canal to be predictive of faecal incontinence ⩾grade 1, and lateral-posterior subregions near the ano-rectal junction to be associated with rectal bleeding ⩾grade 2. Where endpoints are comparable, the SRRs determined in the present study generally, but not exclusively, encompass these regions of interest. The observed variation in SRR clusters and patterns supports the notion that discrete toxicity effects may have different pathophysiologies.

A novelty of the work presented here is not only the comparison of SRR and full RW dose-toxicity analysis for planned dose, but also for total accumulated dose. The rectum is highly susceptible to motion and deformation [31], and therefore planned dose is not equal to delivered dose [8,9,14,16,41]. When comparing planned and delivered dose, total accumulated daily dose to the RW was systematically lower than planned dose ($p < 0.001$). This effect has been reported previously [14,16] and is thought to be due to interfraction motion causing a blurring of the highest dose regions. However, when focusing on SRRs, accumulated dose was greater than planned dose ($p < 0.001$), an effect masked when considering the full RW. This may indicate that these SRRs are more greatly affected by interfraction motion, and therefore dose calculated at treatment planning is not representative of the dose actually received in these regions. Differences between planned and delivered dose limit the level of accuracy achievable in dose-toxicity modelling.

In lieu of widely available systems for accumulating daily delivered dose to the rectum, previous studies have addressed this problem using approaches such as estimating motion-inclusive dose through statistical simulation [42] or by extrapolation of dose calculated using a sample of cone-beam CT (CBCT) scans acquired throughout the course of treatment [14,41]. Casares-Magaz et al. conducted a case-control study of 24 patients whereby accumulated dose was estimated through rigid registration of dose from 13 CBCT scans per patient. Dose-surface histograms of RW DSMs were analysed for associations with gastrointestinal toxicity ⩾grade 2. Results revealed a subregion in the inferior rectum from the accumulated DSM predictive of toxicity, where planned dose could not be considered predictive. This supports the hypothesis that delivered dose is a better predictor of rectal toxicity than planned dose.

For accumulated dose, the use of SRRs produced stronger associations with toxicity than the full RW for three of four endpoints; rectal bleeding, proctitis, and faecal incontinence. For rectal bleeding this suggested radiosensitive regions spanning the inferior circumference to the right-posterior mid-RW, and an anterior-superior cluster. For proctitis, this was focussed in the posterior to lateral inferior RW. Small SRR clusters in the mid-posterior and right-inferior RW were identified for faecal incontinence. Results for stool frequency could not be considered significant, nor could any result from planned dose, other than EUD to the full RW for rectal bleeding. Of the three endpoints found to be associated with EUD to SRRs, although accumulated AUCs were greater than from planned dose, there was considerable overlap of CIs for proctitis and faecal incontinence, suggesting no significant advantage to using accumulated dose over planned dose. ORs from accumulated SRRs were greater than from planned SRRs for rectal bleeding and incontinence. All other results had CIs crossing 1. The greatest improvement to dose-toxicity association when analysing accumulated dose to SRRs was found for rectal bleeding, where the AUC was 13% greater than planned dose to the full RW and 15% greater than planned EUD to SRR, with a corresponding increase in OR of 26%.

The issue of correcting for multiple comparison testing in voxelwise dose-toxicity studies has been debated in the scientific literature. Palma et al. [43] promote correcting for family-wise error rate (such as Bonferroni correction) or false discovery rate (such as the Benjamini-Hochberg, as investigated here). However, others have suggested that p-value adjustment may lead to increasing sensitivity to the point of overfitting the data when applied to voxel-level DSM analysis [10]. In their analysis of bladder DSMs, Palorini et al. [44] discussed the issue at depth, and argued against reducing or adjusting p-values in this context, emphasising that their main goal was 'to identify the shape of dose regions that show the highest discrimination between the two groups of patients (toxicity versus no toxicity) without looking at significance at the single-pixel level'. They applied a similar post-processing smoothing approach to that used in this study. Based on these recommendations, Onjukka et al. [10] reported results for unadjusted p-values in a recent pattern analysis of rectal DSMs. Results for adjusted p-values were also included, but were considered overly restrictive. In the present study, results based on adjusted p-values have been included for completeness, but the focus was on SRRs generated using unadjusted p-values.

Higher AUCs were expected for accumulated dose than planned dose as, intuitively, true delivered dose incorporating interfraction motion should be more predictive of toxicity than the planned dose based on static anatomy. However, differences were generally small and there was overlap of CIs. The significance of results based on these small differences, for the relatively low toxicity rates, can therefore be highly sensitive to other limiting factors. For example, here, the accuracy of the FE model is dependent on the rectal contours, and accumulated dose data can be limited by the MVCT field of view.

The power of the current study was limited by its sample size. Although similar to previous studies [10,18], here this resulted in relatively large confidence intervals, particularly when considering validation results with respect to training. The number of validation data points was lower than that generally considered adequate for ROC/AUC analysis, so OR comparisons were also included to provide further interpretation of results. Analysis of the validation set could not be considered to verify the training set results. However, these results should be interpreted with caution. The poor performance and wide confidence intervals are likely due to the small sample size and event rate within the validation set. The rationale for using a data partitioning approach (rather than alternative internal validation methods such as bootstrapping or k-fold cross-validation) was to ensure that the testing of the SRR analysis was conducted using an independent dataset, kept blind from the SRR-generation process. This was the approach used by Dréan et al. [18] in their work determining rectal SRRs using a similar sample size.

The assumption was made that the two fractionation regimes could be radiobiologically combined due to equivalent levels of toxicity incidence [45]. This approach has been used in similar investigations of rectal toxicity [10,18]. However, within the patient cohort, increased rates of faecal incontinence and stool frequency were observed in the 74 Gy arm. Conversely, in a recent study by Heemsbergen et al. [15], increased rates of toxicity were observed for rectal bleeding and stool frequency in a hypofractionated arm (64 Gy in 3.4 Gy fractions) compared with the conventional fractionation (78 Gy in 2 Gy fractions). The current study does not consider patient co-factors (including prescription), but such effects may be investigated in the future when a larger patient cohort is available.

The incorporation of accumulated dose to SRRs into dose-toxicity analysis has been shown to be more beneficial than using planned dose alone for certain toxicity endpoints, particularly rectal bleeding. The inclusion of accumulated dose into toxicity prediction models becomes an increasingly important consideration with the move towards hypo and ultra-hypofractionated treatments such as prostate SABR, where daily interfraction motion has a much greater effect in terms of contribution towards total delivered dose. Additionally, the duration of these treatment deliveries is of the order of magnitude where intrafraction motion should also be considered [46]. The effect of intrafraction motion was outwith the scope of the VoxTox research programme.

In conclusion, we have demonstrated that voxel-level dose accumulation using FE-DSMs facilitates higher resolution interrogation of associations between rectal SRRs and toxicity. Regions of significant dose differences were identified per endpoint, away from the highest dose levels. SRRs defined from accumulated dose were more predictive

of rectal bleeding than from planned dose. However, validation results were not significant. Ultimately, the aim is to use voxel-level accumulated dose information to develop more accurate and robust toxicity prediction models than previously achievable from planned dose. In the context of adaptive radiotherapy, these models could be used to identify and monitor patients most at risk of developing toxicity. If required, the treatment could then be modified to minimise and control the risk of toxicity, with a view to reducing incidence rates, lowering medical costs of associated care, and improving the patient's quality of life.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank the patients who participated in the study, the referring physicians, and VoxTox Research Facilitator, Michael Simmons. We would like to acknowledge Dr. Agnieszka Lemanska, University of Surrey, who performed external validation of toxicity mapping rules. LEAS was supported by the University of Cambridge W D Armstrong Trust Fund; AMB, MR and KH were supported by the VoxTox Programme Grant, which was funded by Cancer Research UK (CRUK), ref: C8857/A13405; DJN was supported by a CRUK Clinical Research Fellowship; NGB is supported by the National Institute for Health Research (NIHR) Manchester Biomedical Research Centre.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.phro.2020.05.006.

References

[1] Emami B, Lyman J, Brown A, Cola L, Goitein M, Munzenrider J, et al. Tolerance of normal tissue to therapeutic irradiation. Int J Radiat Oncol Biol Phys 1991;21(1):109–22. https://doi.org/10.1016/0360-3016(91)90171-Y.
[2] Marks LB, Yorke ED, Jackson A, Ten Haken RK, Constine LS, Eisbruch A, et al. Use of normal tissue complication probability models in the clinic. Int J Radiat Oncol Biol Phys 2010;76(3 Suppl)(1):S10–9.https://doi.org/10.1016/j.ijrobp.2009.07.1754.
[3] Barnett GC, West CML, Dunning AM, Elliott RM, Coles CE, Pharoah PDP, et al. Normal tissue reactions to radiotherapy: towards tailoring treatment dose by genotype. Nat Rev Cancer 2009;9(2):134–42. https://doi.org/10.1038/nrc2587.
[4] Jaffray DA, Lindsay PE, Brock KK, Deasy JO, Tomé W. Accurate accumulation of dose for improved understanding of radiation effects in normal tissue. Int J Radiat Oncol Biol Phys 2010;76(3):S135–9. https://doi.org/10.1016/j.ijrobp.2009.06.093.
[5] Wortel RC, Witte MG, van der Heide UA, Pos FJ, Lebesque JV, van Herk M, et al. Dose–surface maps identifying local dose–effects for acute gastrointestinal toxicity after radiotherapy for prostate cancer. Radiother Oncol 2015;117(3):515–20. https://doi.org/10.1016/j.radonc.2015.10.020.
[6] Acosta O, Drean G, Ospina JD, Simon A, Haigron P, Lafond C, et al. Voxel-based population analysis for correlating local dose and rectal toxicity in prostate cancer radiotherapy. Phys Med Biol 2013;58(8):2581. https://doi.org/10.1088/0031-9155/58/8/2581.
[7] Buettner F, Gulliford S, Partridge M, Sydes MR, Dearnaley D, Webb S. Assessing correlations between the spatial distribution of dose to the rectal wall and late rectal toxicity after prostate radiotherapy. IFMBE Proc 2009;25(1):908–11. https://doi.org/10.1007/978-3-642-03474-9-255.
[8] Murray J, McQuaid D, Dunlop A, Buettner F, Nill S, Hall E, et al. SU-E-J-14: a novel approach to evaluate the dosimetric effect of rectal variation during image guided prostate radiotherapy. Med Phys 2014;41(6):157. https://doi.org/10.1118/1.4888065.
[9] Scaife JE, Thomas SJ, Harrison K, Romanchikova M, Sutcliffe MPF, Forman JR, et al. Accumulated dose to the rectum, measured using dose-volume histograms and dose-surface maps, is different from planned dose in all patients treated with radiotherapy for prostate cancer. Brit J Radiol 2015;88(1054):20150243. https://doi.org/10.1259/bjr.20150243.
[10] Onjukka E, Fiorino C, Cicchetti A, Palorini F, Improta I, Gagliardi G, et al. Patterns in ano-rectal dose maps and the risk of late toxicity after prostate IMRT. Acta Oncol:1–8.https://doi.org/10.1080/0284186X.2019.1635267.
[11] Moulton CR, House MJ, Lye V, Tang CI, Krawiec M, Joseph DJ, et al. Spatial features of dose–surface maps from deformably-registered plans correlate with late gastrointestinal complications. Phys Med Biol 2017;62.
[12] Casares-Magaz O, Muren LP, Moiseenko V, Petersen SE, Pettersson NJ, Høyer M, et al. Spatial rectal dose/volume metrics predict patient-reported gastro-intestinal symptoms after radiotherapy for prostate cancer. Acta Oncol 2017;56(11):1507–13. https://doi.org/10.1080/0284186X.2017.1370130.
[13] Vanneste BG, Buettner F, Pinkawa M, Lambin P, Hoffmann AL. Ano-rectal wall dose-surface maps localize the dosimetric benefit of hydrogel rectum spacers in prostate cancer radiotherapy. Clin Transl Rad Oncol 2018;14:17–24. https://doi.org/10.1016/j.ctro.2018.10.006.
[14] Casares-Magaz O, Bülow S, Pettersson NJ, Moiseenko V, Pedersen J, Thor M, et al. High accumulated doses to the inferior rectum are associated with late gastro-intestinal toxicity in a case-control study of prostate cancer patients treated with radiotherapy 2019;58(10):1543–1546.https://doi.org/10.1080/0284186X.2019.1632476.
[15] Heemsbergen WD, Incrocci L, Pos FJ, Heijmen BJ, Witte MG. Local dose effects for late gastrointestinal toxicity after hypofractionated and conventionally fractionated modern radiotherapy for prostate cancer in the HYPRO trial. Front Oncol 2020;10:469. https://doi.org/10.3389/fonc.2020.00469.
[16] Shelley LEA, Scaife JE, Romanchikova M, Harrison K, Forman JR, Bates AM, et al. Delivered dose can be a better predictor of rectal toxicity than planned dose in prostate radiotherapy. Radiother Oncol 2017;123:466–71. https://doi.org/10.1016/j.radonc.2017.04.008.
[17] Dréan G, Acosta O, Lafond C, Simon A, de Crevoisier R, Haigron P. Inter-individual registration and dose mapping for voxelwise population analysis of rectal toxicity in prostate cancer radiotherapy. Med Phys 2016;43(6):2721–30. https://doi.org/10.1118/1.4948501.
[18] Dréan G, Acosta O, Ospina JD, Fargeas A, Lafond C, Corrégé G, et al. Identification of a rectal subregion highly predictive of rectal bleeding in prostate cancer IMRT. Radiother Oncol 2016;119(3):388–97. https://doi.org/10.1016/j.radonc.2016.04.023.
[19] Brock K, Sharpe M, Dawson L, Kim S, Jaffray D. Accuracy of finite element model-based multi-organ deformable image registration. Med Phys 2005;32(6):1647–59. https://doi.org/10.1118/1.1915012.
[20] Brock KK, Dawson LA, Sharpe MB, Moseley DJ, Jaffray DA. Feasibility of a novel deformable image registration technique to facilitate classification, targeting, and monitoring of tumor and normal tissue. Int J Radiat Oncol Biol Phys 2006;64(4):1245–54. https://doi.org/10.1016/j.ijrobp.2005.10.027.
[21] Velec M, Moseley JL, Craig T, Dawson LA, Brock KK. Accumulated dose in liver stereotactic body radiotherapy: positioning, breathing, and deformation effects. Int J Radiat Oncol Biol Phys 2012;83(4):1132–40. https://doi.org/10.1016/j.ijrobp.2011.09.045.
[22] Boubaker MB, Haboussi M, Ganghoffer JF, Aletti P. Finite element simulation of interactions between pelvic organs: predictive model of the prostate motion in the context of radiotherapy. J Biomech 2009;42(12):1862–8. https://doi.org/10.1016/j.jbiomech.2009.05.022.
[23] Boubaker MB, Haboussi M, Ganghoffer JF, Aletti P. Predictive model of the prostate motion in the context of radiotherapy: a biomechanical approach relying on urodynamic data and mechanical testing. J Mech Behav Biomed 2015;49:30–42. https://doi.org/10.1016/j.jmbbm.2015.04.016.
[24] Velec M, Moseley JL, Svensson S, Hårdemark B, Jaffray DA, Brock KK. Validation of biomechanical deformable image registration in the abdomen, thorax, and pelvis in a commercial radiotherapy treatment planning system. Med Phys 2017;44(7):3407–17. https://doi.org/10.1002/mp.12307.
[25] Zambrano V, Furtado H, Fabri D, LUtgendorf-Caucig C, GOra J, Stock M, et al. Performance validation of deformable image registration in the pelvic region. J Radiat Res 2013;54(suppl 1):i120–i128.https://doi.org/10.1093/jrr/rrt045.
[26] Brock KK. Image processing in radiation therapy. CRC Press; 2014.
[27] Burnet NG, Scaife J, Romanchikova M, Thomas S, Bates A, Wong E, et al. Applying physical science techniques and CERN technology to an unsolved problem in radiation treatment for cancer: the multidisciplinary 'VoxTox' research programme. CERN IdeaSquare J Exp Innov 2017;1(1):3.
[28] Dearnaley D, Syndikus I, Sumo G, Bidmead AM, Bloomfield D, Clark C, et al. Conventional versus hypofractionated high-dose intensity-modulated radiotherapy for prostate cancer: preliminary safety results from the CHHiP randomised controlled trial. Lancet Oncol 2012;13(1):43–54. https://doi.org/10.1016/S1470-2045(11)70293-5.
[29] Dearnaley D, Syndikus I, Mossop H, Khoo V, Birtle A, Bloomfield D, et al. Conventional versus hypofractionated high-dose intensity-modulated radiotherapy for prostate cancer: 5-year outcomes of the randomised, non-inferiority, phase 3 CHHiP trial. Lancet Oncol 2016;17(8):1047–60. https://doi.org/10.1016/S1470-2045(16)30102-4.
[30] Burnet N, Adams E, Fairfoul J, Tudor G, Hoole A, Routsis D, et al. Practical aspects of implementation of helical tomotherapy for intensity-modulated and image-guided radiotherapy. Clin Oncol 2010;22(4):294–312. https://doi.org/10.1016/j.clon.2010.02.003.
[31] Scaife J, Harrison K, Romanchikova M, Parker A, Sutcliffe M, Bond S, et al. Random variation in rectal position during radiotherapy for prostate cancer is two to three times greater than that predicted from prostate motion. Brit J Radiol 2014;87(1042):20140343. https://doi.org/10.1259/bjr.20140343.
[32] Shelley LEA, Sutcliffe MPF, Harrison K, Scaife JE, Parker MA, Romanchikova M, et al. Autosegmentation of the rectum on megavoltage image guidance scans. BioMed Phys Eng Express 2019;5(2):025006https://doi.org/10.1088/2057-1976/aaf1db.
[33] Scaife JE, Harrison K, Drew A, Cai X, Lee J, Schonlieb CB, et al. Accuracy of manual and automated rectal contours using helical tomotherapy image guidance scans

during prostate radiotherapy. J Clin Oncol 2015;33(7_suppl). https://doi.org/10.1200/jco.2015.33.7_suppl.94. 94–94.
[34] Thomas SJ, Romanchikova M, Harrison K, Parker MA, Bates AM, Scaife JE, et al. Recalculation of dose for each fraction of treatment on TomoTherapy. Brit. J Radiol 2016;89(1059). https://doi.org/10.1259/bjr.20150770.
[35] Thomas S, Eyre KR, Tudor GSJ, Fairfoul J. Dose calculation software for helical tomotherapy, utilizing patient CT data to calculate an independent three-dimensional dose cube. Med Phys 2012;39(1):160–7. https://doi.org/10.1118/1.3668061.
[36] Wilkins A, Naismith OF, Brand D, Fernandez K, Hall E, Dearnaley DP, et al. Derivation of dose/volume constraints for the anorectum from clinician-reported outcomes (CRO) in the CHHiP trial of radiotherapy (RT) fractionation. J Clin Oncol 2018;36(6):87. https://doi.org/10.1200/jco.2018.36.6_suppl.87.
[37] U.S. department of health and human services NIoH. National Cancer Institute. Common Terminology Criteria for Adverse Events (CTCAE) Version 4.0 2010. http://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03_2010-06-14_QuickReference_5x7.pdf. accessed: July 2016 (2010).
[38] LENT SOMA tables table of contents. Radiother Oncol 1995;35(1):17–60.https://doi.org/10.1016/0167-8140(95)90055-1.
[39] Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Bhttps://doi.org/10.1111/j.2517-6161.1995.tb02031.x.
[40] Niemierko A. Reporting and analyzing dose distributions: a concept of equivalent uniform dose. Med Phys 1997;24(1):103–10. https://doi.org/10.1118/1.598063.
[41] Thor M, Bentzen L, Hysing LB, Ekanger C, Helle SI, Karlsdóttir Á, et al. Prediction of rectum and bladder morbidity following radiotherapy of prostate cancer based on motion-inclusive dose distributions. Radiother Oncol 2013;107(2):147–52. https://doi.org/10.1016/j.radonc.2013.03.029.
[42] Thor M, Apte A, Deasy JO, Muren LP. Statistical simulations to estimate motion-inclusive dose-volume histograms for prediction of rectal morbidity following radiotherapy. Acta Oncol 2013;52(3):666–75. https://doi.org/10.3109/0284186X.2012.720382.
[43] Palma G, Monti S, Cella L. Voxel-based analysis in radiation oncology: a methodological cookbook. Physica Medica 2020;69:192–204. https://doi.org/10.1016/j.ejmp.2019.12.013.
[44] Palorini F, Cozzarini C, Gianolini S, Botti A, Carillo V, Iotti C, et al. First application of a pixel-wise analysis on bladder dose–surface maps in prostate cancer radiotherapy. Radiother Oncol 2016;119(1):123–8. https://doi.org/10.1016/j.radonc.2016.02.025.
[45] Dearnaley D, Syndikus I, Mossop H, Khoo V, Birtle A, Bloomfield D, et al. Conventional versus hypofractionated high-dose intensity-modulated radiotherapy for prostate cancer: 5-year outcomes of the randomised, non-inferiority, phase 3 chhip trial. Lancet Oncol 2016;17(8):1047–60. https://doi.org/10.1016/S1470-2045(16)30102-4.
[46] Thomas SJ, Ashburner M, Tudor GSJ, Treeby J, Dean J, Routsis D, et al. Intra-fraction motion of the prostate during treatment with helical tomotherapy. Radiother Oncol 2013;109(3):482–6. https://doi.org/10.1016/j.radonc.2013.09.011.