Relationships between rectal and perirectal doses and rectal bleeding or tenesmus in pooled voxel-wise analysis of 3 randomised phase 3 trials

Keywords: External beam radiotherapy (EBRT), prostate cancer, gastrointestinal toxicity, rectal bleeding, tenesmus, voxel-wise analysis, dose-toxicity relationships

M Marcello MSc, BSc [a, b]

JW Denham MD, FRANZCR [c]

A Kennedy BSc(Hons) [b]

A Haworth FACPSEM, PhD, MSc, BSc(Hons) [d]

A Steigler BMath [e]

PB Greer PhD, MSc, BSc [f, g]

LC Holloway PhD, BSc [h, i, j]

JA Dowling PhD, BComp(Hons I), BAppSc [f, k]

MG Jameson PhD, B.Med.Rad.PhysSc(Hons) [h, i, j, l]

D Roach MSc, BS [h, i, l]

DJ Joseph MD, FRANZCR, MRACMA [m, n, o]

SL Gulliford PhD, MSc(Hons), BSc(Hons) [p, q]

DP Dearnaley MA, MD, FRCR, FRCP [r]

MR Sydes MSc, CStat [s]

E Hall PhD, BSc [t]

MA Ebert PhD, BSc(Hons) [a, b, n]

[a] Department of Physics, University of Western Australia, Western Australia, Australia
(Address: 35 Stirling Highway, Crawley, WA, 6009, Australia)

[b] Department of Radiation Oncology, Sir Charles Gairdner Hospital, Western Australia, Australia
(Address: Hospital Avenue, Nedlands, WA, 6009, Australia)

[c] School of Medicine and Public Health, University of Newcastle, New South Wales, Australia
(Address: University Drive, Callaghan, NSW, 2308, Australia)

[d] School of Physics, University of Sydney, New South Wales, Australia
(Address: Physics Road, Camperdown, NSW, 2006, Australia)

[e] Prostate Cancer Trials Group, School of Medicine and Public Health, University of Newcastle, New South Wales, Australia
(Address: University Drive, Callaghan, NSW, 2308, Australia)

[f] School of Mathematical and Physical Sciences, University of Newcastle, New South Wales, Australia
(Address: University Drive, Callaghan, NSW, 2308, Australia)

[g] Department of Radiation Oncology, Calvary Mater Newcastle, New South Wales, Australia
(Address: Corner of Edith & Platt Street, Waratah, NSW, 2298, Australia)

[h] Department of Medical Physics, Liverpool Cancer Centre, New South Wales, Australia
(Address: 1 Campbell Street, Liverpool, NSW, 2170, Australia)

[i] South Western Sydney Clinical School, University of New South Wales, New South Wales, Australia
(Address: Goulburn Street, Liverpool, NSW, 2170, Australia)

[j] Centre for Medical Radiation Physics, University of Wollongong, New South Wales, Australia
(Address: Northfields Avenue, Wollongong, NSW, 2522, Australia)

[k] CSIRO, Queensland, Australia
(Address: Butterfield Street, Herston QLD 4029, Australia)

[l] Cancer Research Team, Ingham Institute for Applied Medical Research, New South Wales, Australia
(Address: 1 Campbell Street, Liverpool, NSW, 2170, Australia)

[m] School of Surgery, University of Western Australia, Western Australia, Australia
(Address: 35 Stirling Highway, Crawley, WA, 6009, Australia)

[n] 5D Clinics, Claremont, Western Australia, Australia
(Address: 261 Stirling Highway, Claremont, WA, 6010, Australia)

[o] GenesisCare WA, Western Australia, Australia
(Address: 24 Salvado Road, Wembley, WA, 6014, Australia)

[p] Radiotherapy Department, University College London Hospitals NHS Foundation Trust, London, United Kingdom
(Address: Gower Street, Bloomsbury, London, WC1E 6BT, United Kingdom)

[q] Department of Medical Physics and Biomedical Engineering, University College London, London, UK
(Address: Gower Street, Bloomsbury, London, WC1E 6BT, United Kingdom)

[r] Academic UroOncology Unit, The Institute of Cancer Research and the Royal Marsden NHS Trust, London, United Kingdom
(Address: 15 Cotswold Road, Sutton, London, SM2 5NG, Australia)

[s] MRC Clinical Trials Unit, Medical Research Council, United Kingdom
(Address: 90 High Holborn, Holborn, London WC1V 6LJ, Australia)

[t] Clinical Trials and Statistics Unit, The Institute of Cancer Research, London, United Kingdom
(Address: Physics Rd, Camperdown, NSW, 2006, Australia)

Corresponding author details:

Mr Marco Marcello
Department of Physics
University of Western Australia
Tel: 0439 940 621
Email: 20739859@student.uwa.edu.au

DATA SHARING STATEMENT

The data from this study was derived from three clinical trials (the RADAR, RT01 and CHHiP trials). The authors do not own these data and hence are not permitted to share them in the original form (only in aggregate form, e.g., publications).

ABSTRACT

Background and purpose: This study aimed to identify anatomically-localised regions where planned radiotherapy dose is associated with gastrointestinal toxicities in healthy tissues throughout the pelvic anatomy.

Materials and methods: Planned dose distributions for up to 657 participants of the Trans-Tasman Radiation Oncology Group 03.04 RADAR trial were deformably registered onto a single exemplar CT dataset. Voxel-wise multiple comparison permutation dose-difference testing, Cox regression modelling and LASSO feature selection were used to identify regions where dose-increase was associated with grade≥2 rectal bleeding (RB) or tenesmus, according to the LENT/SOMA scale. This was externally validated by registering dose distributions from the RT01 (n=388) and CHHiP (n=241) trials onto the same exemplar and repeating the tests on each of these data sets, and on all three datasets combined.

Results: Voxel-wise Cox regression and permutation dose-difference testing revealed regions where increased dose was correlated with gastrointestinal toxicity. Grade≥2 RB was associated with posteriorly extended lateral beams that manifested high doses (> 55 Gy) in a small rectal volume adjacent to the CTV. A correlation was found between grade≥2 tenesmus and increased low-intermediate dose (~25 Gy) at the posterior beam region, including the posterior rectum and perirectal fat space (PRFS).

Conclusions: The serial response of the rectum with respect to RB has been demonstrated in participants with posteriorly extended lateral beams. Similarly, the parallel response of the PRFS with respect to tenesmus has been demonstrated in participants treated with the posterior beam.

INTRODUCTION

Gastrointestinal (GI) symptoms remain a commonly-reported side-effect of prostate external beam radiotherapy (EBRT)[1,2], despite improvements in precision from new technologies such as image guided intensity modulated radiotherapy (IG-IMRT)[3]. Improving the accuracy of predictive toxicity models, through further understanding of dose-volume toxicity relationships, will help optimise organ at risk (OAR) dose constraints. Most current models provide constraints based on information from dose-volume histograms (DVHs) describing planned dose to whole OARs[4,5]. This ignores heterogeneous intra-organ radio-sensitivity, which can be investigated through spatial dose information in the 3D planned distribution not utilised by whole-organ DVHs. Voxel-wise analyses, in which anatomically localised dose-toxicity relationships are identified on the surface or within the volume of OARs, can determine radiosensitive subregions of OARs from which more optimal dose constraints may be derived.

Several studies have begun to investigate rectal dose-sensitivity in this manner. Buettner et al investigated the relationship between late rectal toxicities and spatial features from rectal dose-surface maps[6]. Rectal bleeding (RB) and loose stools were shown to be more strongly correlated with these features than rectal dose-surface histograms (DSH). This group proceeded to parameterise the 3D dose distribution to the rectum and correlate resulting features with late RB, loose stools, and a global toxicity score[7]. These features predicted all three endpoints more accurately than standard DVHs. In a similar analysis by Moulton et al, spatial features of rectal dose-surface maps related to dose shape and coverage were shown to correlate strongly with a range of late GI complications[8]. Other voxel-wise studies have identified predictive rectal subregions[9,10]. DVHs derived from these subregions (found within the inferior-anterior rectum) were demonstrated to be more predictive than whole-rectum DVHs.

No study to date, however, has performed a voxel-wise analysis searching for correlation between variation in intended dose within individual voxels and GI toxicity throughout the entire pelvic anatomy. While still able to determine anatomically localised dose-toxicity relationships, this will test the assumption that these relationships are limited to within OAR volumes or surfaces. This may provide understanding of how broader dose patterns relate to toxicity, revealing the impact of factors related to treatment technique. It may also illuminate dose-response pathology that defies assumptions of responsible anatomy.

In this study, multiple voxel-wise statistical methods were employed to investigate the association between 3D planned dose and measures of GI toxicity in the entire pelvic anatomy. Many shortcomings have typically hindered recent voxel-wise analyses[11,12], including misregistration of planned 3D dose distributions ("dosemaps"), false positive rates due to the large number of voxels being statistically compared, not using time-to-event data, or not controlling for patient baseline characteristics. This study used a combination of statistical approaches to compensate these shortcomings. High quality planned dose data from three prospective multi-centre prostate radiotherapy clinical trials was utilised in order to assess the consistency of derived associations across cohorts, participating centres, employed radiotherapy techniques and overall treatment approach.

METHODS AND MATERIALS

Clinical Trials

Table 1 describes the three clinical trials from which data was sourced for this study. 3D planned dose distributions, with corresponding CT images including delineated CTV, rectum and bladder collected by the RADAR trial, were utilised as the primary dataset of this study. Similar information was collected in the RT01 and CHHiP trials and utilised as external validation datasets.

3D Data Preparation

Three CT image registration templates were chosen from an independent cohort of 39 prostate EBRT patients[13]. Pairwise registrations of CT images within this cohort along with registrations between this cohort and the RADAR CT dataset were used to generate a normalised cross correlation similarity matrix. This matrix was used to perform clustering by affinity propagation to select the single most representative patient CT as an exemplar from the initial cohort. This exemplar was the first registration template (T1). Next, an anti-exemplar, most-different from T1, was chosen as a template on which the impact of registration and reference geometry could be tested (T2). Finally, a similar process was used to select a cropped exemplar, enabling analysis to be restricted to a small region including the prostate and immediate surrounding organs (T3). Dose distributions were then deformed onto these templates through application of deformation vector fields obtained from the image-based registrations above. See appendix section 2 for images of templates and registration pipelines. The 3D dose distributions from all phases of radiotherapy were summed together according to biologically isoeffective 2 Gy per fraction dose (EQD2)[14], using a spatially invariant alpha/beta ratio of 3, resulting in a single distribution for each participant registered onto each template. All subsequent analyses used dose distributions which uniformly sampled 1 in 2 voxels for T1 and T2 (due to the large number of total voxels). For T3, every voxel was used.

Gastrointestinal Toxicity Outcomes

Two time-to-event GI toxicity outcome measures were included for analysis: rectal bleeding and tenesmus. An event consisted of the first peak grade≥2 occurrence during follow-up. All participants who reported baseline symptoms of grade≥1 were removed from analysis, apart from RT01 tenesmus participants as this information was not available. Physician assisted toxicity grading was performed

according to the Late Effects on Normal Tissue, Subjective, Objective, Management, Analytic (LENT/SOMA) questionnaire[15]. For RADAR, participants were routinely followed up, post-treatment, every 3 months to 18 months, every 6 months to 5 years, and then annually. RT01 participants were assessed at 6, 12, 18, and 24 months after commencing radiotherapy, and annually thereafter. CHHiP participants were assessed for late toxicities beginning 26 weeks after the start of radiotherapy and every 6 months for 2 years, and then yearly.

Note that all voxel-wise tests were repeated for both outcome measures, on all three trial datasets (RADAR, RT01 and CHHiP), as well as on a dataset combining participants from all trials ("Combined"). All three registration templates were used for RADAR for exploration of dose-toxicity associations, but only T1 for RT01, CHHiP and Combined for validation. The permutation and uni-voxel tests were performed using MATLAB R2016b and later versions (MathWorks, Natick MA), while the multi-voxel LASSO test was performed on R 3.6.1 (The R Foundation, Vienna). All 3D results were displayed using ITK-SNAP version 3.8.0[16].

Voxel-Wise Permutation Dose-Difference Testing

Voxel-wise permutation dose-difference testing was performed according to the method outlined by Chen et al[11]. Following Figure 1, for each given outcome participants were divided according to whether they experienced a toxicity outcome event at any time during follow-up. The mean dose distributions of each group were then compared to each other, voxel-by-voxel, to reveal regions of statistically significant dose-difference. This method utilises a nonparametric permutation-based test in which the group labels are randomly swapped (permuted) and the dose-comparison repeated for each permutation. In this study, 1000 permutations were performed (as recommended by Chen et al) generating a distribution of test statistics which provided a threshold to determine the region of dose-difference with statistical certainty - accounting for the multiple testing problem arising from comparing a vast number of voxels (see Appendix A of Chen et al for more detail). The dose-difference region is produced by thresholding at any chosen p-value. In this study, thresholds of $p<0.05$, $p<0.1$, $p<0.2$ and $p<0.3$ were applied. The test was performed in both directions, determining regions where participants with toxicity have significantly more dose than participants without toxicity, and vice versa.

Uni-Voxel Cox Regression Test

This method generates a separate Cox proportional hazards model for each voxel (hence, 'uni'-voxel), testing the dose-toxicity association in that voxel. Taking a given voxel (out of a total of 664,000 for T1), participants were divided into two groups about the median of the combined distribution of dose values, as in Figure 1. The hazard ratio (HR) of toxicity between the high dose value group and low dose value group was then calculated, with corresponding p-value testing whether $HR>1$ or $HR<1$ with statistical significance. This HR therefore compares the incidence of toxicity between each dose group, indicating the dose- toxicity relationship at the given voxel. Age, prescribed dose, disease risk, cancer stage, baseline PSA concentration and number of treatment beams were patient baseline characteristics investigated as potential control variables in each model, attempting to eliminate their confounding influence at each voxel[17,18], chosen through an automated selection process (see appendix section 1 for details). Repeating this process for every voxel produced a 3D HR map and corresponding p-value map revealing the dose-toxicity across the pelvic anatomy.

Multi-Voxel Cox Regression Test with LASSO Feature Selection

In contrast to the uni-voxel regression, this method combined all voxels in the pelvic anatomy as variables in a single multivariate Cox regression model (hence, 'multi'-voxel). The LASSO (Least Absolute Shrinkage Selection Operator[19]) was then applied to select voxels that did not correlate with each other in the model, while still correlating strongly with the toxicity outcome. The LASSO requires a pre-specified variable, λ, that determines the threshold by which features or variables

(voxels) in the Cox model are selected. As λ increases, more features are excluded, until none are selected. 100 values of λ were pre-specified, equally spaced from that which selected all voxels to that which selected none. For each value of λ, one-in-ten cross validation was used to test the predictive ability of the resulting Cox model – the model comprised of the voxels selected by the LASSO. The final value of λ was that which maximised the corresponding model's ability to predict the toxicity outcome by minimising the partial likelihood deviance. 3D maps were then produced and imposed on the registration templates, imposing the selected voxels on the pelvic anatomy, indicating whether HR>1 or HR<1 in each case. HRs still compared the incidence of toxicity between the high dose group and low dose group at a given voxel, with the cut-point for dose determined in the same way as the uni-voxel test. The LASSO enables selection of voxels strongly correlated with the given outcome while accounting for inter-voxel dose correlation and the multiple testing problem.

RESULTS

Table 2 shows the number of participants from each trial included in the analyses, with corresponding patient baseline characteristic and outcome information, after participants were excluded due to loss of follow-up, missing data, and considering only participants receiving EBRT alone. For RADAR, 6 and 11 early (≤3 months) toxicity events were included for RB and tenesmus respectively. All other toxicity events for all other datasets were late (>3 months).

The tests identified voxel clusters (VCs) and individual voxels across the pelvic anatomy where dose variation was associated with both GI toxicity outcome measures.

Figure 2 shows the results for RB. The dominant pattern is an association between increased RB and posteriorly extended lateral beam margins culminating in higher dose at the rectum adjacent to the CTV centre. Dose-difference maps from all trial cohorts and the combined cohort ("Combined") exhibit this pattern. In particular, the axial planes show RB participants having between approximately 4 and 7 Gy more dose on average than non-RB participants where lateral beams extend posteriorly into the rectal space. The corresponding sagittal planes reveal RB participants have between 4 and 5 Gy more dose on average in a sub-volume of the rectum adjacent to the CTV, statistically different at $p<0.3$ for Combined. Figure 4 shows the mean and standard deviation (SD) dosemaps for the Combined cohort. A voxel was selected in the identified sub-volume adjacent to the CTV to compare doses of participants with and without rectal bleeding here. Figure 4 reveals that Combined participants with and without rectal bleeding had an average dose of 57.3 Gy and 53.0 Gy in this representative voxel respectively (SD=14.2 Gy). The corresponding uni-voxel HR maps confirm these patterns for all datasets, revealing VCs with HR>1 in the same regions, with $p<0.05$ for all cohorts but CHHiP. For CHHiP, VCs with HR>1 ($p<0.05$) were found directly posterior to the rectum. For Combined, the multi-voxel LASSO selected voxels in regions corresponding to where the uni-voxel test revealed associations, with one HR>1 voxel selected at the rectum adjacent to the CTV, visible in the sagittal plane. Increased RB is associated with reduced dose in the anterior beam region for Combined, with HR<1 ($p<0.05$) VCs found here by both the permutation and uni-voxel tests. Less prominently, a correlation is present between increased RB and increased dose in the posterior oblique beam regions for RT01, revealed in the corresponding uni-voxel HR map.

The results for tenesmus are displayed in Figure 3. The most consistent pattern is an association between increased tenesmus and increased dose at the posterior rectum extending posteriorly where the posterior beam is expected to contribute dose. This is clearly seen in the RADAR, CHHiP and Combined results. The permutation test identified VCs of significant dose-difference in this posterior region, where participants with tenesmus have up to 4 Gy more dose for RADAR ($p<0.3$) and Combined ($p<0.05$), with this region being larger and less fragmented for Combined. Figure 4a)

shows that Combined participants with and without tenesmus had an average dose of 24.2 Gy and 21.1 Gy at a voxel in this VC directly posterior to the rectum. The same figure shows the dose-effect region is broad – extending across approximately two thirds of the rectum in the superior-inferior direction and to the surface from the rectum posterior. The uni-voxel HR maps reveal VCs with HR>1 ($p<0.05$) in this same region for RADAR, CHHiP and Combined. For RT01, the same correlation pattern is present, but is weaker, with HR>1 ($p>0.05$) VCs present in this posterior region. The multi-voxel LASSO confirmed this association for Combined by selecting several voxels with HR>1 in the space postero-inferior to the rectum, most visible in the coronal plane. Reduced dose in the anterior beam region near the surface is associated with increased tenesmus for Combined, where participants with tenesmus have up to 3.5 Gy less dose ($p<0.05$), confirmed by the corresponding uni-voxel HR map. Reduced dose is also associated with tenesmus in the oblique and lateral beam regions for Combined, and in the posterior obliques only for RADAR.

DISCUSSION

In this study, quality-assured and reviewed planning data collected in multi-centre clinical trials with extensive follow-up was used to derive independent datasets for analysis. Subsequent associations between voxel-dose and measures of GI toxicity across the pelvic anatomy have been identified without assuming associations necessarily occur at organ sites. Although no individual voxel-wise test in this study addressed every typical shortcoming of voxel-wise analyses, each test did address particular shortcomings such that a consistent result across all techniques could be considered independent of these issues.

Rectal bleeding was consistently correlated with increased dose in rectal sub-volumes adjacent to the CTV centre, manifesting in participants treated with posteriorly extended lateral beams. This sub-volume is located at the boundary of the high-dose region. Moving directly posteriorly from the CTV, this sub-volume coincides almost exactly with the maximum standard deviation in dose in the combined cohort along this plane. Therefore, this association is most likely not indicating that this rectal sub-volume is particularly radiosensitive relative to the rest of the rectum. It is more likely that a dose-bleeding effect is highlighted within this particular sub-volume because it is located in both the high-dose region and in a region of sufficient dose variation to reveal this statistical effect.

All three trials allowed 70 Gy to 15% of the rectal volume. Participants with posteriorly extended lateral beams may have plans that met dose-volume constraints while still resulting in high doses (up to 70 Gy) to small volumes (up to 15%) of the rectum, resulting in RB. This comports with studies that have established correlations between high doses to small volumes of the rectum and RB across multiple treatment modalities[29,30,31,32,33]. 98.1% of grade≥2 RB events in the combined cohort were late (>3 months). Late radiation damage includes progressive obliterative endarteritis that leads to ischaemia and fibrosis of the rectal tissue, ulcerating and eroding rectal blood vessels, resulting in bleeding[34]. This study provides the first 3D evidence of the serial response of the rectum with respect to late rectal bleeding without the assumption that dose-toxicity effects necessarily occur at organ sites, while relating this effect to treatment technique. It is therefore recommended that posterior extension of lateral beams be done with an awareness of the potential to produce rectal hotspots associated with RB.

In contrast to RB, tenesmus was correlated with increased dose in the posterior beam region and decreased dose in the posterior oblique beam regions. Correlation coincided with the posterior perirectal fat space (PRFS). The dose-effect here was distributed across a broad volume and was in the low-intermediate dose-range (~25Gy). Therefore, the effect may be due to increased low-intermediate doses broadly distributed in the PRFS. Ebert et al have shown a dose-tenesmus response at the anal canal and anorectum in the low-intermediate dose range (5-38 Gy)[29]. Moulton et al have

similarly demonstrated an association between tenesmus and greater low-intermediate doses (20-30 Gy) to the inferior 20% and lateral-posterior of the rectum in combined EBRT and high dose-rate brachytherapy participants[8]. An association between increased low-intermediate doses (10-40 Gy) throughout the PRFS and grade≥2 tenesmus was demonstrated by Gulliford et al, suggesting the PRFS responds as a parallel structure with respect to control related toxicities[35]. These findings are broadly consistent with this study, noting that all three studies utilised participants from the RADAR trial. The PRFS facilitates rectal motility, compliance and control, and contains a large number of sympathetic, parasympathetic and non-autonomic nerve fibres[35]. Damage to this region may therefore lead to nerve dysfunction, contributing to control related symptoms such as tenesmus. The broad posterior beam association found here confirms the PRFS behaving as a parallel structure with respect to tenesmus. The posterior beam should be used with an awareness of this behaviour.

This study has utilised planned dose distributions. These will differ from delivered dose distributions in practice[36]. It has been shown that delivered dose can be a better predictor of rectal toxicity than planned dose[37]. As the consistency between planned and delivered dose improves, or delivered dose becomes increasingly measurable, voxel-wise dose analyses will become more effective in finding anatomically localised dose-outcome relationships. Data derived from participants treated with IGRT, for example, would ensure planned dose more closely resembles delivered dose. The accuracy of registration and the appropriateness of the choice of exemplar and anti-exemplar could impact derived results. A perfectly accurate registration would ensure associations are in fact occurring at corresponding anatomical sites. Diversity in the dose distributions across the cohort is also limiting, as the mean dosemaps are approximate 3 or 4 field treatments in all datasets (see appendix section 5). Greater diversity in technique will enable more generalisable feature selection. Only the uni-voxel Cox regression accounted for participant baseline factors, and these represent only a sample of possible factors that could confound the identified dose-toxicity relationships. To ensure these relationships are independent of a given baseline factor, separating the cohort into this factor's subgroups prior to analysis is necessary. This, however, would reduce power, requiring a larger cohort. It must also be noted that participants from the RT01 cohort had the longest follow-up time, while those from CHHiP had the shortest. This means the likelihood of experiencing the considered toxicities would be different for all three trials and this effect was not controlled for. The longest possible follow-up times were considered to maximise power.

This was the first study performing a full voxel-wise analysis of dose-rectal toxicity relationships in the entire pelvic anatomy. Previous studies have established a relationship between high doses in small rectal volumes and resulting bleeding, and subsequent studies have determined predictive rectal subregions[38,10]. This study has reinforced this rectal dose-bleeding relationship, further substantiating the work done in moving toward defining constraints based on subregions as opposed to whole-rectum DVHs. This study has also uniquely identified broader dose-rectal toxicity patterns. Namely that the use posteriorly extended lateral beams may produce rectal dose hotspots not prevented by conventional dose constraints, resulting in increased incidence of late RB. Also, the use of the posterior beam can lead to higher low-intermediate doses in the PRFS, increasing risk of tenesmus and potentially other control related symptoms.

ACKNOWLEDGEMENTS

We acknowledge funding from the Australian National Health and Medical Research Council (grants 300705, 455521, 1006447, 1077788), the Hunter Medical Research Institute, the Health Research Council (New Zealand), the University of Newcastle, the Calvary Mater Newcastle, the Medical Research Council Clinical Trials Unit at University College London, Abbott Laboratories and Novartis Pharmaceuticals. We acknowledge funding form the Medical Research Council UK (grant MC_UU_12023/28) for the RT01 trial. David Dearnaley, Emma Hall and Sarah Gulliford

acknowledge NHS funding to the NIHR Biomedical Research Centre at the Royal Marsden NHS Foundation Trust and The Institute of Cancer Research, London. We acknowledge all trial investigators and participants who've made this study possible. We gratefully acknowledge the support of the Sir Charles Gairdner Hospital, Rachel Kearvell, the 'Elvis' study team including Kristie Harrison, participating RADAR centres, the Trans-Tasman Radiation Oncology Group, Ben Hooton and Elizabeth van der Wath. We are also grateful for the contributions of Oscar Acosta, Renaud de Crevoisier, and Eugenia Mylona.

REFERENCES

1. Zelefsky, M. J. *et al.* Incidence of late rectal and urinary toxicities after three-dimensional conformal radiotherapy and intensity-modulated radiotherapy for localized prostate cancer. *Int. J. Radiat. Oncol. Biol. Phys.* **70**, 1124–1129 (2008).

2. Denham, J. W. *et al.* Rectal and urinary dysfunction in the TROG 03.04 RADAR trial for locally advanced prostate cancer. *Radiother. Oncol.* **105**, 184–192 (2012).

3. Wortel, R. C. *et al.* Acute toxicity after image-guided intensity modulated radiation therapy compared to 3D conformal radiation therapy in prostate cancer patients. *Int. J. Radiat. Oncol. Biol. Phys.* **91**, (2015).

4. Cambria, R. *et al.* Evaluation of Late Rectal Toxicity after Conformal Radiotherapy for Prostate Cancer A Comparison between Dose-Volume Constraints and NTCP Use. *Strahlenther Onkol* **185**, 384–390 (2009).

5. Marks, L. B. *et al.* Use of normal tissue complication probability models in the clinic. *Int. J. Radiat. Oncol. Biol. Phys.* **76**, S10–S19 (2010).

6. Buettner, F. *et al.* Assessing correlations between the spatial distribution of the dose to the rectal wall and late rectal toxicity after prostate radiotherapy: an analysis of data from the MRC RT01 trial (ISRCTN 47772397). *Phys. Med. Biol.* **54**, 6535–48 (2009).

7. Buettner, F., Gulliford, S. L., Webb, S. & Partridge, M. Modeling late rectal toxicities based on a parameterized representation of the 3D dose distribution. *Phys. Med. Biol.* **56**, 2103–2118 (2011).

8. Moulton, C. R. *et al.* Spatial features of dose – surface maps from deformably-registered plans correlate with late gastrointestinal complications. *Phys. Med. Biol.* **62**, 4118–4139 (2017).

9. Acosta, O. *et al.* Voxel-based population analysis for correlating local dose and rectal toxicity in prostate cancer radiotherapy. *Phys. Med. Biol.* **58**, 2581–2595 (2013).

10. Dréan, G. *et al.* Identification of a rectal subregion highly predictive of rectal bleeding in prostate cancer IMRT. *Radiother. Oncol.* **119**, 388–397 (2016).

11. Chen, C., Witte, M., Heemsbergen, W. & van Herk, M. Multiple comparisons permutation test for image based data mining in radiotherapy. *Radiat. Oncol.* **8**, 293 (2013).

12. Ospina, J. D. *et al.* Spatial nonparametric mixed-effects model with spatial-varying coefficients for analysis of populations. in *International Workshop on Machine Learning in Medical Imaging* 142–150 (Springer, 2011).

13. Kennedy, A. *et al.* Similarity clustering based atlas selection for pelvic CT image segmentation. *Med. Phys.* **46**, 2246–2250 (2019).

14. Bentzen, S. M. *et al.* Bioeffect modeling and equieffective dose concepts in radiation oncology-Terminology, quantities and units. *Radiother. Oncol.* **105**, 266–268 (2012).

15. Pavi, J., Denekamp, J. & Letschert, J. LENT-SOMA scales for all anatomic sites. *Int J Radiat Oncol Biol Phys* **31**, 1049–1091 (1995).

16. Yushkevich, P. A. *et al.* User-guided 3D active contour segmentation of anatomical structures:

Significantly improved efficiency and reliability. *Neuroimage* **31**, 1116–1128 (2006).

17. VanderWeele, T. J. & Shpitser, I. On the definition of a confounder. *Ann. Stat.* **41**, 196–220 (2013).

18. Marcello, M. *et al.* Association between treatment planning and delivery factors and disease progression in prostate cancer radiotherapy: Results from the TROG 03.04 RADAR trial. *Radiother. Oncol.* **126**, 249–256 (2018).

19. Tibshiranit, B. R. Regression Shrinkage and Selection via the Lasso. *J. R. Stat. Soc. Ser. B* **58**, 267–288 (1996).

20. Ebert, M. A. *et al.* Gastrointestinal Dose-Histogram Effects in the Context of Dose-Volume Constrained Prostate Radiation Therapy : Analysis of Data From the RADAR Prostate Radiation Therapy Trial. *Int. J. Radiat. Oncol. Biol. Phys.* **91**, 595–603 (2014).

21. Dréan, G. *et al.* Identification of a rectal subregion highly predictive of rectal bleeding in prostate cancer IMRT. *Radiother. Oncol.* **119**, 388–397 (2016).

22. Colaco, R. J. *et al.* Rectal Toxicity After Proton Therapy For Prostate Cancer : An Analysis of Outcomes of Prospective Studies Conducted at the University of Florida Proton Therapy Institute. *Radiat. Oncol. Biol.* **91**, 172–181 (2014).

23. Fukahori, M., Matsufuji, N., Himukai, T., Kanematsu, N. & Mizuno, H. Estimation of late rectal normal tissue complication probability parameters in carbon ion therapy for prostate cancer. *Radiother. Oncol.* **118**, 136–140 (2016).

24. Moulton, C. R. *et al.* Prostate external beam radiotherapy combined with high-dose-rate brachytherapy: dose-volume parameters from deformably-registered plans correlate with late gastrointestinal complications. *Radiat. Oncol.* 1–13 (2016). doi:10.1186/s13014-016-0719-2

25. Reis, E. D., Vine, A. J. & Heimann, T. Radiation damage to the rectum and anus: Pathophysiology, clinical features and surgical implications. *Color. Dis.* **4**, 2–12 (2002).

26. Gulliford, S. L. *et al.* Radiotherapy dose-distribution to the perirectal fat space (PRS) is related to gastrointestinal control-related complications. *Clin. Transl. Radiat. Oncol.* **7**, 62–70 (2017).

27. Colvill, E. *et al.* Multileaf Collimator Tracking Improves Dose Delivery for Prostate Cancer Radiation Therapy : Results of the First Clinical Trial. *Radiat. Oncol. Biol.* **92**, 1141–1147 (2015).

28. Shelley, L. E. A. *et al.* Delivered dose can be a better predictor of rectal toxicity than planned dose in prostate radiotherapy. *Radiother. Oncol.* **123**, 466–471 (2017).

29. Acosta, O. *et al.* Voxel-based population analysis for correlating local dose and rectal toxicity in prostate cancer radiotherapy. *Phys. Med. Biol.* **58**, 2581–95 (2013).

30. Denham, J. W. *et al.* Short-term androgen suppression and radiotherapy versus intermediate-term androgen suppression and radiotherapy, with or without zoledronic acid, in men with locally advanced prostate cancer (TROG 03.04 RADAR): An open-label, randomised, phase 3 factorial. *Lancet Oncol.* **15**, 1076–1089 (2014).

31. Denham, J. W. *et al.* Radiation dose escalation or longer androgen suppression for locally advanced prostate cancer? Data from the TROG 03.04 RADAR trial. *Radiother. Oncol.* **115**, 301–307 (2015).

32. Dearnaley, D. P. *et al.* Escalated-dose versus standard-dose conformal radiotherapy in prostate cancer : first results from the MRC RT01 randomised controlled trial. *Lancet Oncol.* **8**, 475–87 (2007).

33. Dearnaley, D. P. *et al.* Escalated-dose versus control-dose conformal radiotherapy for prostate cancer : long-term results from the MRC RT01 randomised controlled trial. *Lancet Oncol.* **15**, 464–473 (2014).

34. Dearnaley, D. *et al.* Conventional versus hypofractionated high-dose intensity-modulated radiotherapy for prostate cancer: preliminary safety results from the CHHiP randomised controlled trial. *Lancet Oncol.* **13**, 43–54 (2012).

35. Dearnaley, 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.* **0**, 1840–1850 (2016).

36. Ebert, M. A. *et al.* Detailed review and analysis of complex radiotherapy clinical trial planning data: Evaluation and initial experience with the SWAN software system. *Radiother. Oncol.* **86**, 200–210 (2008).

37. Sydes, M. R. *et al.* Implementing the UK Medical Research Council (MRC) RT01 trial (ISRCTN 47772397): methods and practicalities of a randomised controlled trial of conformal radiotherapy in men with localised prostate cancer. *Radiother. Oncol.* **72**, 199–211 (2004).

38. Naismith, O. *et al.* Radiotherapy Quality Assurance for the CHHiP Trial: Conventional Versus Hypofractionated High-Dose Intensity-Modulated Radiotherapy in Prostate Cancer. *Clin. Oncol.* **31**, 611–620 (2019).

FIGURE CAPTIONS

Figure 1 *Visual representation of the a) Voxel-Wise Dose-Difference Permutation Test, b) Uni-Voxel Cox Regression test and c) Multi-Voxel Cox Regression Test with LASSO Feature Selection.*

Figure 2 *Results for rectal bleeding. Corresponding axial, coronal and sagittal slices (top to bottom) of a) mean dose-difference maps and significant dose-difference regions determined by permutation test, b) uni-voxel Cox regression HR and p-value maps and c) multi-voxel Cox regression LASSO HR maps (with uni-voxel p-values for comparison), for respective data sets. Planes are shown coinciding with significant results, indicated in other reconstructions with dashed lines. Warm colours (yellow, orange, red) indicate increased doses are associated with incidence of the outcome measure, while colder colours (tones of blues)indicate reduced doses are associated with incidence of the outcome measure. The bladder, CTV and rectum are delineated in purple, pink and light blue respectively.*

Figure 3 *Results for tenesmus. Corresponding axial, coronal and sagittal slices (top to bottom) of a) mean dose-difference maps and significant dose-difference regions determined by permutation test, b) uni-voxel Cox regression HR and p-value maps and c) multi-voxel Cox regression LASSO HR maps (with uni-voxel p-values for comparison), for respective data sets. Planes are shown coinciding with significant results, indicated in other reconstructions with dashed lines. Warm colours (yellow, orange, red) indicate increased doses are associated with incidence of the outcome measure, while colder colours (tones of blues)indicate reduced doses are associated with incidence of the outcome measure. The bladder, CTV and rectum are delineated in purple, pink and light blue respectively.*

Figure 4 *a) Map of the mean planned dose for the combined cohort. Displayed on the map are the permutation test dose-difference regions for both RB and tenesmus, with the mean doses and dose-differences at a key voxel within those regions highlighted for each outcome measure. b) Map of the standard deviation planned dose for the combined cohort. Similarly displayed are the permutation test result regions, with the standard deviation dose at a voxel at the centre of the rectal bleeding region.*

Table 1 *Clinical trials information.*

	RADAR	**RT01**	**CHHiP**
Full name	Randomised Androgen Deprivation and Radiotherapy (TROG 03.04) Trial[30,31]	A Randomised Trial of High Dose Therapy in Localised Cancer of the Prostate using Conformal Radiotherapy Techniques[32,33]	Conventional or Hypofractionated High Dose Intensity Modulated Radiotherapy for Prostate Cancer Trial[34,35]
Descriptors	• Randomised • Phase 3 • Factorial	• Randomised • Phase 3 • Superiority	• Randomised • Phase 3 • Non-inferiority
Goal	Comparison of 6 months of androgen deprivation therapy (ADT) plus radiotherapy with 18 months of ADT with the same radiotherapy	Comparison of 64 Gy standard-dose and 74 Gy dose-escalated conformal radiotherapy	Comparison of conventional and hypofractionated IMRT
Countries	Australia and New Zealand	United Kingdom, New Zealand, Australia	United Kingdom, New Zealand, Rep. of Ireland, Switzerland
Accrual years	Oct 2003 – Aug 2007	Jan 1998 – Dec 2001	Oct 2002 – Jun 2011
Total accrued subjects	1071	843	3216
Date data was frozen	June 2015	Aug 2013	Oct 2017
Participants	Intermediate-risk (T2a) or high-risk (T2b+) prostate cancer	T1b – T3a prostate cancer	T1b – T3a prostate cancer
Radiotherapy type	Dose escalated 3D conformal EBRT	Standard or dose escalated 3D conformal EBRT	Dose escalated IMRT
Prescribed dose groups (dose per fraction)	66 Gy (2 Gy), 70 Gy (2 Gy), 74 Gy (2 Gy)	64 Gy (2 Gy), 74 Gy (2 Gy)	57 Gy (3 Gy), 60 Gy (3 Gy), 74 Gy (2 Gy)
Rectal dose-volume constraints	Maximum of 65 Gy, 70 Gy and 75 Gy to 40%, 30% and 5% of rectal volume respectively	A maximum of 64 Gy and 74 Gy to any volume of the rectum for each dose group respectively	Maximum of 65 Gy, 70 Gy and 75 Gy to 30%, 15% and 3% of rectal volume respectively
Beam arrangements	Any preferred combination of 3 or more conformal beams	3 or 4 beams (anterior/lateral/posterior) for first 64 Gy, with additional 4 or 6 beam boost to 74 Gy	3 or 4 beams (anterior/lateral/posterior) or 5 beams or more if inverse planning utilised
Electronic review of treatment planning data	Full retrospectve review for all subjects[36]	No electronic individual plan review[37]	Full prospective case reviews for the first 2 or 3 subjects at each centre[38]
Manager	TROG Cancer Research, NSW, Australia	Medical Research Clinical Trials Unit, London, UK	Clinical Trials and Statistics Unit, the Institute of Cancer Research, London, UK
Trial registration number	ISRCTN90298520	ISRCTN47772397	ISRCTN97182923
Ethics approval number	Approved by Hunter New England Human Research Ethics Committee Trial ID 03/06/11/3.02	North Thames Multi-centre Research Ethics Committee number MREC/97/2/16	Approved by the London Multi-centre Research Ethics Committee number 04/MRE02/10

Table 2 *The number of subjects in each trial dataset, broken down by endpoint and baseline variables, including follow-up information.*

RADAR				RT01			CHHiP			COMBINED		
		Bleeding (grade≥2)	**Tenesmus (grade≥2)**		**Bleeding (grade≥2)**	**Tenesmus (grade≥2)**		**Bleeding (grade≥2)**	**Tenesmus (grade≥2)**		**Bleeding (grade≥2)**	**Tenesmus (grade≥2)**
Total number of subjects		657	657	**Total number of subjects**	363	388	**Total number of subjects**	235	241	**Total number of subjects**	1225	1286
Events		214 (32.6%)	252 (38.4%)	**Events**	91 (25.1%)	71 (18.3%)	**Events**	18 (7.7%)	12 (5.0%)	**Events**	321 (25.6%)	335 (26.0%)
Follow-up in months (min, max, med, IQR)		(3, 96, 48, 54)	(3, 96, 42, 60)	**Follow-up in months (min, max, med, IQR)**	(6, 147, 72, 62)	(12, 158, 98, 65)	**Follow-up in months (min, max, med, IQR)**	(6, 68, 60, 2)	(6, 68, 60, 2)	**Follow-up in months (min, max, med, IQR)**	(3, 147, 60, 36)	(3, 158, 60, 59)
Variables	**Definitions**			**Definitions**			**Definitions**			**Definitions**		
Age[1]	**Median**	69.4 yrs	69.4 yrs	**Median**	68.0 yrs	67.9 yrs	**Median**	67.3 yrs	67.4 yrs	**Median**	68.5 yrs	68.5 yrs
Prescribed dose	**[66 Gy]**	82	82	**[64 Gy]**	173	204	**[57 Gy]**	72	86	**[66 Gy (RADAR), 64 Gy (RT01), 57 Gy and 60 Gy (CHHiP)]**	373	453
	[70 Gy]	366	366	**[74 Gy]**	190	184	**[60 Gy]**	70	80			
	[74 Gy]	209	209				**[74 Gy]**	64	75	**[70 Gy and 74 Gy (RADAR), 74 Gy (RT01), 74 Gy (CHHiP)]**	882	833
Disease risk	**[GS ≤ 7]**	461	461	**[T1b/c or T2a with (PSA + (GS - 6)*10) < 15]**	98	110	**[T1b/c or T2a with PSA ≤ 10 and GS ≤ 6]**	51	58	**[Lower risk group subjects from each respective dataset]**	617	883
	[GS > 7]	196	196									
				[T1b/c or T2a with (PSA + (GS - 6)*10) ≥ 15 or T2b/T3a]	265	278	**[Any of the following: Stage ≥ T2b, 10 < PSA ≤ 20, GS > 6]**	155	183	**[Higher risk group subjects from each respective dataset]**	638	403
Cancer stage	**[T2]**	473	473	**[≤ T2a (T1b, T1c, T2a)]**	221	235	**[≤ T2a (T1a, T1b, T1c, T2a)]**	174	175	**[Lower cancer stage group subjects from each respective dataset]**	865	885
	[T3/T4]	184	184	**[> T2a (T2b, T3a)]**	142	153	**[> T2a (T2b, T2c, T3a)]**	61	66			
										[Higher cancer stage group subjects from each respective dataset]	390	406
Baseline PSA[1]	**Median**	14.00 ng/ml	14.04 ng/ml	**Median**	13.80 ng/ml	13.80 ng/ml	**Median**	11.70 ng/ml	11.70 ng/ml	**Median**	13.60 ng/ml	13.50 ng/ml
Number of beams	**[3 beams]**	67	67	**[3 beams for phase 1 of treament]**	214	228	**[≤ 4 beams]**	204	210	**[≤ 4 beams (RADAR), 3 beams (RT01), ≤ 4 beams (CHHiP)]**	840	858
	[4 beams]	350	350				**[> 4 beams]**	31	31			
				[4 beams for phase 1 of treament]	149	160						
	[5 beams]	88	88							**[> 4 beams (RADAR), 4 beams (RT01), > 4 beams (CHHiP)]**	415	428
	[6 beams]	93	93									
	[≥ 7 beams]	59	59									

[1]This variable was divided into two approximately equal subgroups split about the median value

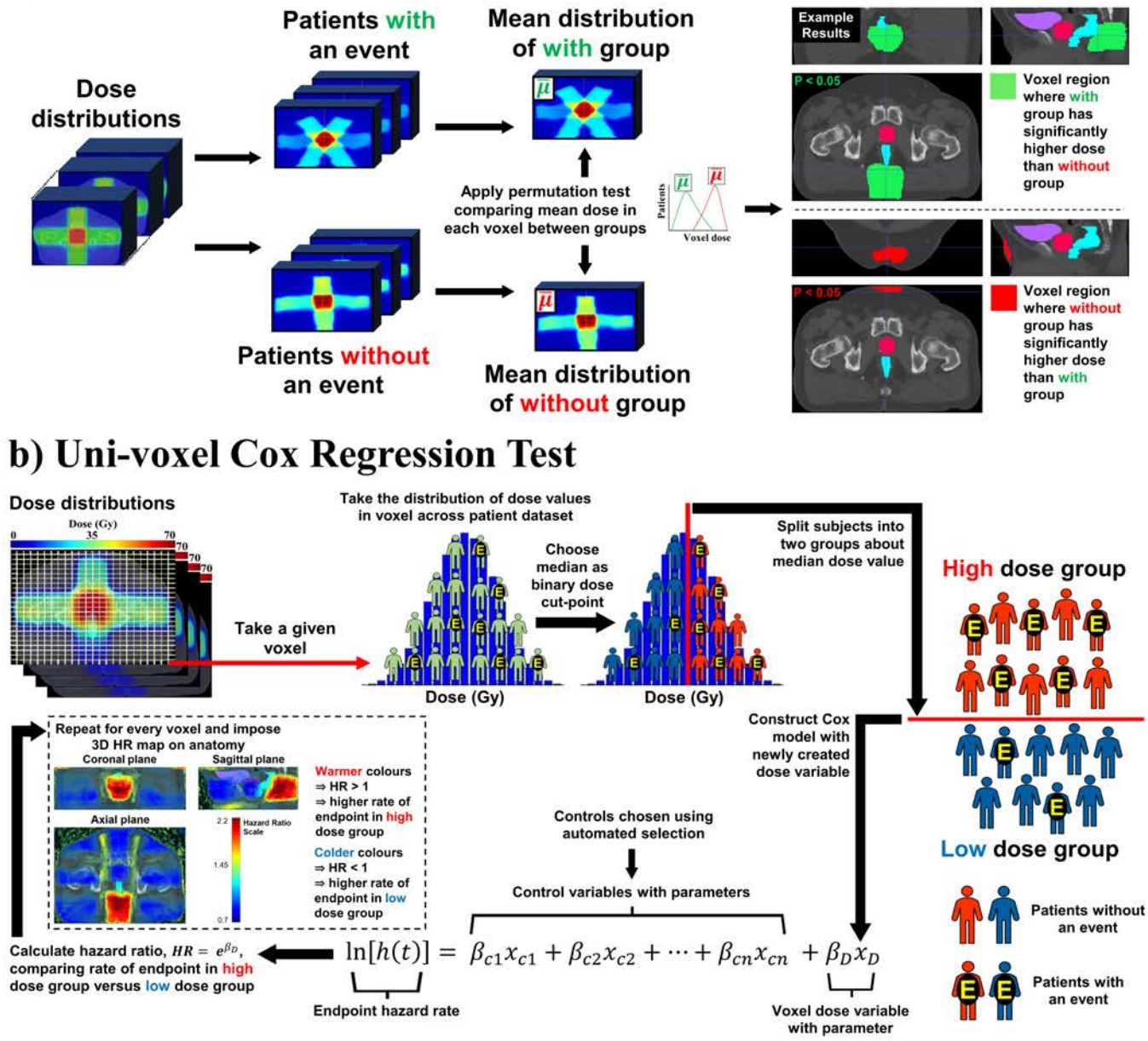
a) Voxel-Wise Dose-Difference Permutation Test
Dose distributions
Patients with an event
Patients without an event
Mean distribution of with group
Mean distribution of without group
Apply permutation test comparing mean dose in each voxel between groups
Patients
Voxel dose
Example Results
P < 0.05
Voxel region where with group has significantly higher dose than without group
Voxel region where without group has significantly higher dose than with group
b) Uni-voxel Cox Regression Test
Dose distributions
Dose (Gy)
Take a given voxel
Take the distribution of dose values in voxel across patient dataset
Choose median as binary dose cut-point
Split subjects into two groups about median dose value
High dose group
Low dose group
Repeat for every voxel and impose 3D HR map on anatomy
Coronal plane
Sagittal plane
Axial plane
Hazard Ratio Scale
Warmer colours ⇒ HR > 1 ⇒ higher rate of endpoint in high dose group
Colder colours ⇒ HR < 1 ⇒ higher rate of endpoint in low dose group
Controls chosen using automated selection
Control variables with parameters
Construct Cox model with newly created dose variable
Calculate hazard ratio, $HR = e^{\beta_D}$, comparing rate of endpoint in high dose group versus low dose group
$\ln[h(t)] = \beta_{c1}x_{c1} + \beta_{c2}x_{c2} + \cdots + \beta_{cn}x_{cn} + \beta_D x_D$
Endpoint hazard rate
Voxel dose variable with parameter
Patients without an event
Patients with an event
c) Multi-Voxel Cox Regression Test with LASSO Feature Selection
Dose distributions
Voxel dose variables for every voxel combined into one Cox proportional hazards model
(as opposed to a separate model for every voxel as in the uni-voxel test)
$\ln[h(t)] = \beta_{D1}x_{D1} + \beta_{D2}x_{D2} + \beta_{D3}x_{D3} + \beta_{D4}x_{D4} + \cdots + \beta_{Dk}x_{Dk}$
Endpoint hazard rate
Voxel dose variables with parameters
Each voxel dose variable was generated in the same way as the uni-voxel analysis
Coronal plane
Sagittal plane
Axial plane
HR > 1
HR < 1
Display selected voxels as HR > 1 or HR < 1 on anatomy.
Note: the magnitude of the HR is not necessary to depict the magnitude of correlation, as its mere selection implies it is highly correlated with the endpoint.
Take the selected variables (voxels) and calculate their hazard ratios (E.g. $HR_{D1} = e^{\beta_{D1}}$, $HR_{D2} = e^{\beta_{D4}}$)
Variable selection
(to select voxels that are most strongly correlated with the endpoint independent of each other)
Take 9/10 of data
Choose a λ value
Apply LASSO
Repeat by cycling through the remaining data (cross-validation) and select variables whose parameters minimise PLD
Repeat for all 100 values of λ and pass on set of selected parameters that minimise PLD
Compute partial likelihood deviance (PLD) using selected model parameters in Cox model applied to remaining 1/10 of data
$\ln[h(t)] = \beta_{D1}x_{D1} + \beta_{D2}x_{D2} + \beta_{D3}x_{D3} + \beta_{D4}x_{D4} + \cdots + \beta_{Dk}x_{Dk}$

Figure 1

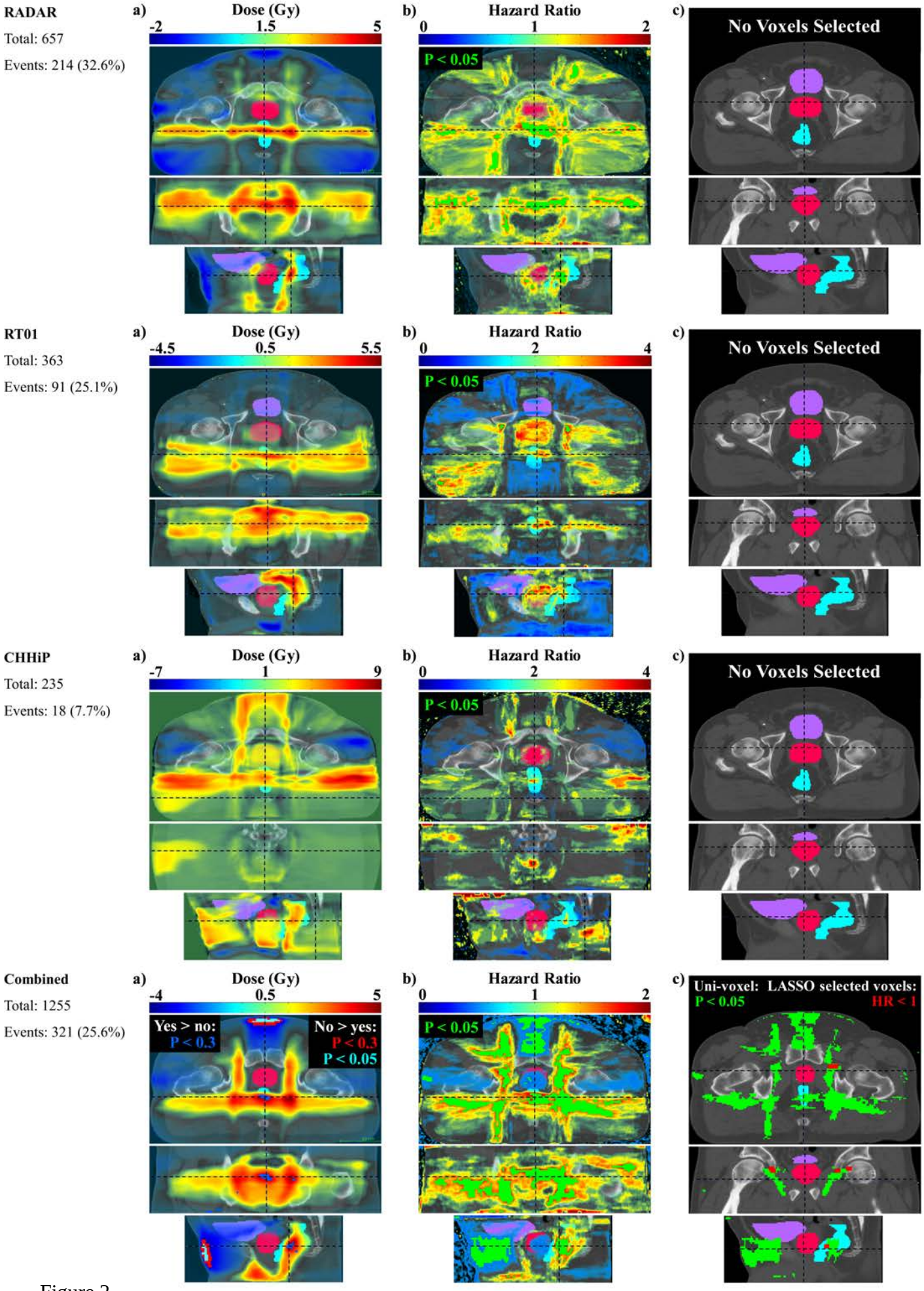
RADAR
Total: 657
Events: 214 (32.6%)
a)
Dose (Gy)
-2
1.5
5
b)
Hazard Ratio
0
1
2
P < 0.05
c)
No Voxels Selected
RT01
Total: 363
Events: 91 (25.1%)
a)
Dose (Gy)
-4.5
0.5
5.5
b)
Hazard Ratio
0
2
4
P < 0.05
c)
No Voxels Selected
CHHiP
Total: 235
Events: 18 (7.7%)
a)
Dose (Gy)
-7
1
9
b)
Hazard Ratio
0
2
4
P < 0.05
c)
No Voxels Selected
Combined
Total: 1255
Events: 321 (25.6%)
a)
Dose (Gy)
-4
0.5
5
Yes > no:
P < 0.3
No > yes:
P < 0.3
P < 0.05
b)
Hazard Ratio
0
1
2
P < 0.05
c)
Uni-voxel: LASSO selected voxels:
P < 0.05
HR < 1

Figure 2

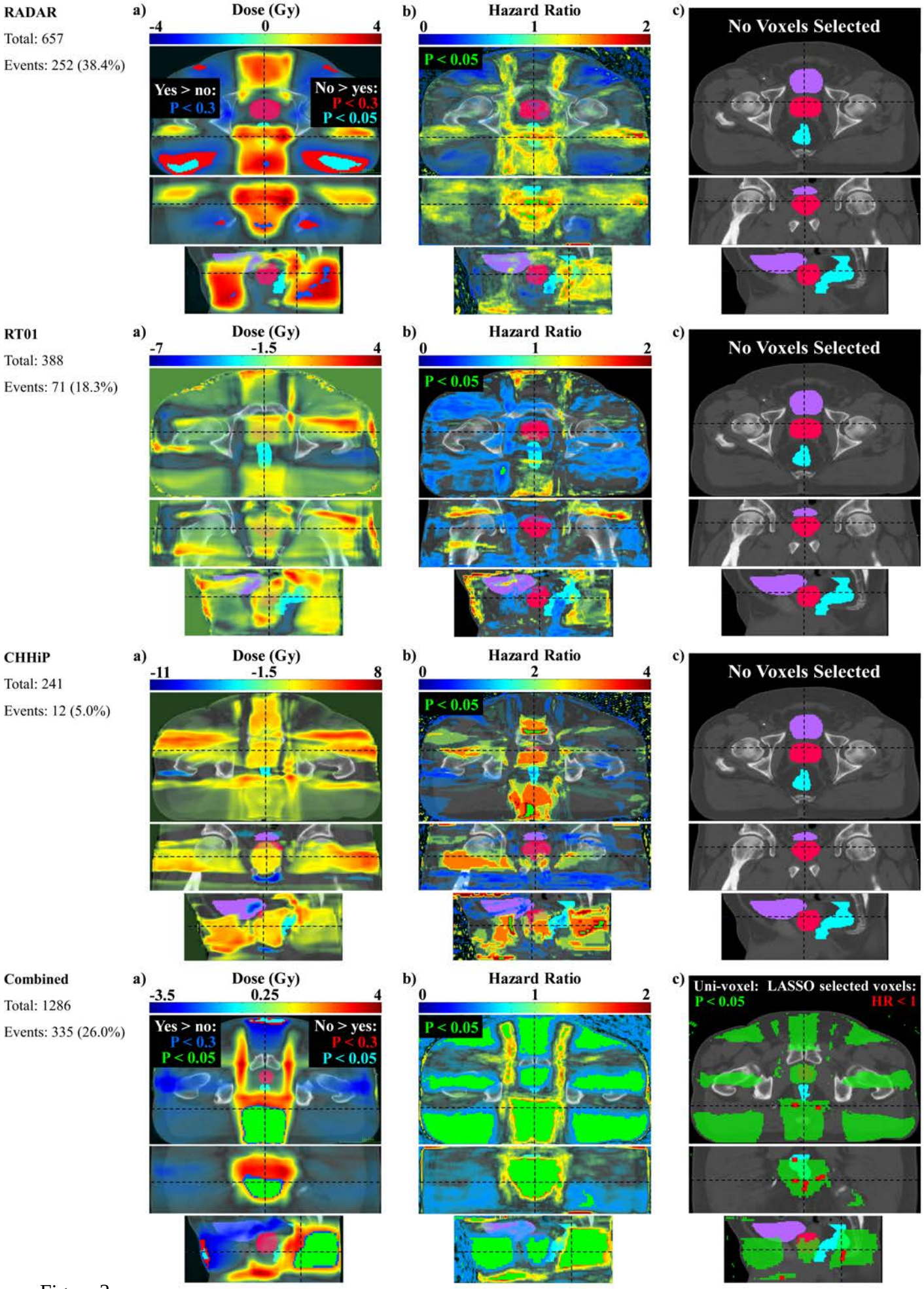
RADAR
Total: 657
Events: 252 (38.4%)
a)
Dose (Gy)
-4
0
4
Yes > no:
P < 0.3
No > yes:
P < 0.3
P < 0.05
b)
Hazard Ratio
0
1
2
P < 0.05
c)
No Voxels Selected
RT01
Total: 388
Events: 71 (18.3%)
a)
Dose (Gy)
-7
-1.5
4
b)
Hazard Ratio
0
1
2
P < 0.05
c)
No Voxels Selected
CHHiP
Total: 241
Events: 12 (5.0%)
a)
Dose (Gy)
-11
-1.5
8
b)
Hazard Ratio
0
2
4
P < 0.05
c)
No Voxels Selected
Combined
Total: 1286
Events: 335 (26.0%)
a)
Dose (Gy)
-3.5
0.25
4
Yes > no:
P < 0.3
P < 0.05
No > yes:
P < 0.3
P < 0.05
b)
Hazard Ratio
0
1
2
P < 0.05
c)
Uni-voxel: LASSO selected voxels:
P < 0.05
HR < 1

Figure 3

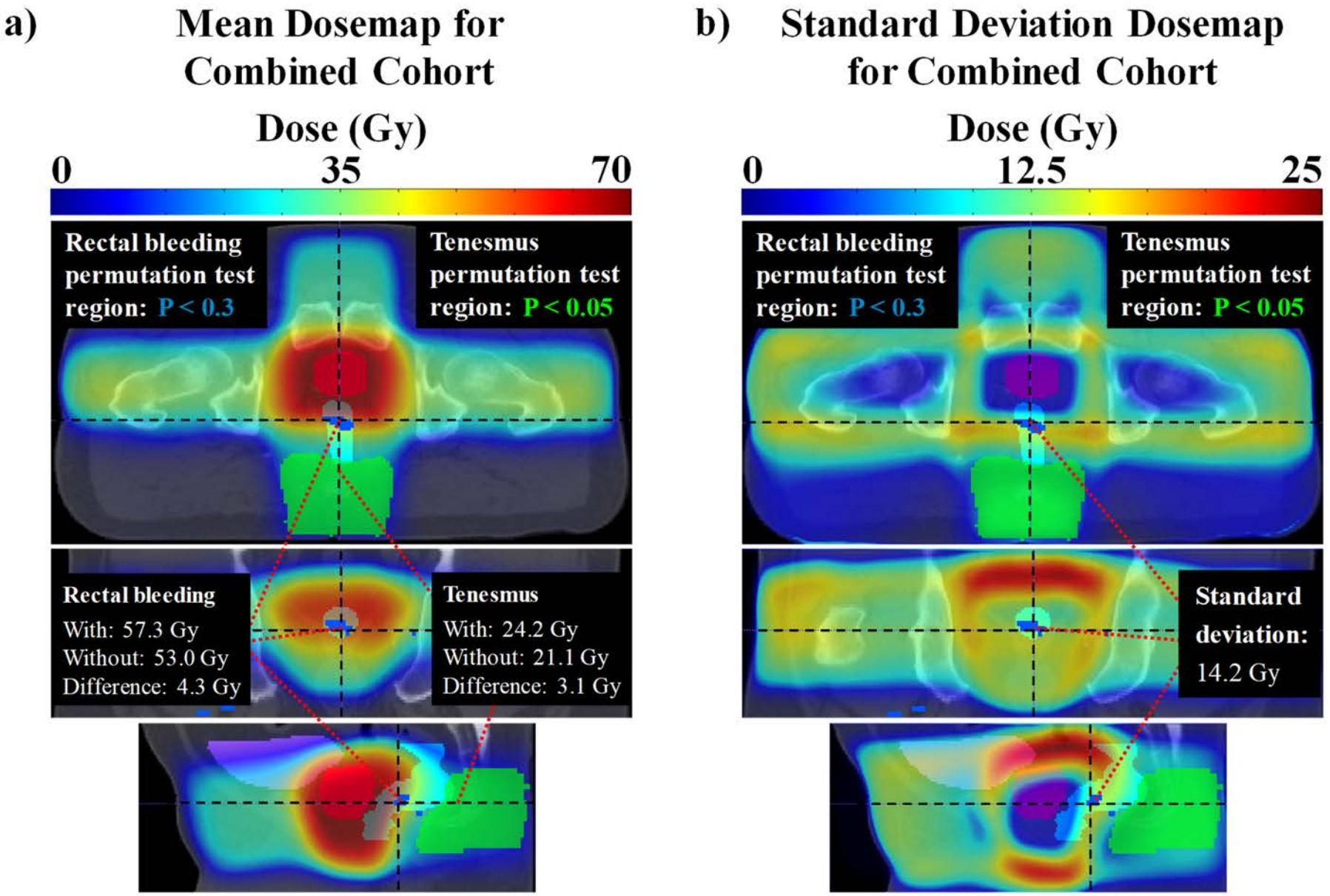
a)
Mean Dosemap for Combined Cohort
Dose (Gy)
0
35
70
Rectal bleeding permutation test region: P < 0.3
Tenesmus permutation test region: P < 0.05
Rectal bleeding
With: 57.3 Gy
Without: 53.0 Gy
Difference: 4.3 Gy
Tenesmus
With: 24.2 Gy
Without: 21.1 Gy
Difference: 3.1 Gy
b)
Standard Deviation Dosemap for Combined Cohort
Dose (Gy)
0
12.5
25
Rectal bleeding permutation test region: P < 0.3
Tenesmus permutation test region: P < 0.05
Standard deviation:
14.2 Gy

Figure 4