

Reduced dose posterior to prostate correlates with increased PSA progression in voxel-based analysis of 3 randomised phase 3 trials

Running title: Voxel-based association of dose and progression

M Marcello MSc, BSc^{1,2}

JW Denham MD, FRANZCR³

A Kennedy BSc(Hons)²

A Haworth FACPSEM, PhD, MSc, BSc(Hons)⁴

A Steigler BMath⁵

PB Greer PhD, MSc, BSc^{6,7}

LC Holloway PhD, BSc^{8,9,10}

JA Dowling PhD, BComp(Hons I), BAppSc^{6,11}

MG Jameson PhD, B.Med.Rad.PhysSc(Hons)^{8,9,10,12}

D Roach MSc, BS^{8,9,12}

DJ Joseph MD, FRANZCR, MRACMA^{13,14,15}

SL Gulliford PhD, MSc(Hons), BSc(Hons)^{16,17}

DP Dearnaley MA, MD, FRCR, FRCP¹⁸

MR Sydes MSc, CStat¹⁹

E Hall PhD, CStat²⁰

MA Ebert PhD, BSc(Hons)^{1,2,14}

¹ Department of Physics, University of Western Australia, Western Australia, Australia
(Address: 35 Stirling Highway, Crawley, WA, 6009, Australia)

² Department of Radiation Oncology, Sir Charles Gairdner Hospital, Western Australia, Australia
(Address: Hospital Avenue, Nedlands, WA, 6009, Australia)

³ School of Medicine and Public Health, University of Newcastle, New South Wales, Australia
(Address: University Drive, Callaghan, NSW, 2308, Australia)

⁴ School of Physics, University of Sydney, New South Wales, Australia
(Address: Physics Road, Camperdown, NSW, 2006, Australia)

⁵ Prostate Cancer Trials Group, School of Medicine and Public Health, University of Newcastle, New South Wales, Australia
(Address: University Drive, Callaghan, NSW, 2308, Australia)

⁶ School of Mathematical and Physical Sciences, University of Newcastle, New South Wales, Australia
(Address: University Drive, Callaghan, NSW, 2308, Australia)

⁷ Department of Radiation Oncology, Calvary Mater Newcastle, New South Wales, Australia
(Address: Corner of Edith & Platt Street, Waratah, NSW, 2298, Australia)

- ⁸ Department of Medical Physics, Liverpool Cancer Centre, New South Wales, Australia
(Address: 1 Campbell Street, Liverpool, NSW, 2170, Australia)
- ⁹ South Western Sydney Clinical School, University of New South Wales, New South Wales, Australia
(Address: Goulburn Street, Liverpool, NSW, 2170, Australia)
- ¹⁰ Centre for Medical Radiation Physics, University of Wollongong, New South Wales, Australia
(Address: Northfields Avenue, Wollongong, NSW, 2522, Australia)
- ¹¹ CSIRO, Queensland, Australia
(Address: Butterfield Street, Herston QLD 4029, Australia)
- ¹² Cancer Research Team, Ingham Institute for Applied Medical Research, New South Wales, Australia
(Address: 1 Campbell Street, Liverpool, NSW, 2170, Australia)
- ¹³ School of Surgery, University of Western Australia, Western Australia, Australia
(Address: 35 Stirling Highway, Crawley, WA, 6009, Australia)
- ¹⁴ 5D Clinics, Claremont, Western Australia, Australia
(Address: 261 Stirling Highway, Claremont, WA, 6010, Australia)
- ¹⁵ GenesisCare WA, Western Australia, Australia
(Address: 24 Salvado Road, Wembley, WA, 6014, Australia)
- ¹⁶ Radiotherapy Department, University College London Hospitals NHS Foundation Trust, London, United Kingdom
(Address: Gower Street, Bloomsbury, London, WC1E 6BT, United Kingdom)
- ¹⁷ Department of Medical Physics and Biomedical Engineering, University College London, London, UK
(Address: Gower Street, Bloomsbury, London, WC1E 6BT, United Kingdom)
- ¹⁸ 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)
- ¹⁹ MRC Clinical Trials Unit, Institute of Clinical Trials and Methodology, University College London, United Kingdom
(Address: 90 High Holborn, London, WC1B 9LJ, United Kingdom)
- ²⁰ Clinical Trials and Statistics Unit, The Institute of Cancer Research, London, United Kingdom
(Address: 15 Cotswold Rd, Sutton, SM2 5NG, United Kingdom)

Corresponding author details:

Mr Marco Marcello

Department of Physics

University of Western Australia

35 Stirling Highway

Crawley, Western Australia 6009

Tel: 0439 940 621

Email: 20739859@student.uwa.edu.au

All statistical analyses were performed by the corresponding author.

CLINICAL TRIAL INFORMATION

If the following code names are interchanged exactly as prescribed, the manuscript, figures and tables should make sense.

‘Trial-A’ = RADAR

‘Trial-A’s name’ = Trans-Tasman Radiation Oncology Group 03.04 RADAR trial

‘Trial-A’s full name’ = Randomised Androgen Deprivation and Radiotherapy

‘Trial-A’s code’ = TROG 03.04

‘Trial-A’s manager’ = Trans-Tasman Radiation Oncology Group (TROG)

‘Trial-B’ = RT01

‘Trial-B’s manager’ = Medical Research Clinical Trials Unit, London, UK

‘Trial-C’ = CHHiP

‘Trial-C’s manager’ = Clinical Trials and Statistics Unit at the Institute of Cancer Research, London, UK

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 from the Medical Research Council UK (grant MC_UU_12023/28) for the MRC RT01 trial. David Dearnaley, Emma Hall and Sarah Gulliford acknowledge NHS funding to the National Institute for Health Research (NIHR) Biomedical Research Centre at the Royal Marsden NHS Foundation Trust and The Institute of Cancer Research, London. We acknowledge support of Cancer Research UK (C8262/A7253, C1491/A9895, C1491/A15955, SP2312/021), the Department of Health, the NIHR Cancer Research Network for the CHHiP trial. 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.

DISCLOSURE STATEMENT

Disclosures: None.

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, eg, publications).

Note: I've included Table 1 here and not in a separate document as it was almost impossible to present a separate blinded version of this table as it contains information concerning the trials all the way through.

Table 1 Clinical trials information. See appendix section 7 for expanded margin and dose definitions for different phases of the CHHiP trial.

	RADAR	RT01	CHHiP
Full name	Randomised Androgen Deprivation and Radiotherapy (TROG 03.04) Trial ^{4,5}	A Randomised Trial of High Dose Therapy in Localised Cancer of the Prostate using Conformal Radiotherapy Techniques ^{6,7}	Conventional or Hypofractionated High Dose Intensity Modulated Radiotherapy for Prostate Cancer Trial ^{8,9}
Descriptors	<ul style="list-style-type: none"> • Randomised • Phase 3 • Factorial 	<ul style="list-style-type: none"> • Randomised • Phase 3 • Superiority 	<ul style="list-style-type: none"> • 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)
Margins for prostate treatment volumes	<p>‘1’ refers to phase 1, ‘2’ refers to phase 2</p> <p>GTV = pros + extra capsular extension (intermediate risk)</p> <p>GTV = pros + SV + extra capsular extension (high risk)</p> <p>CTV = GTV (both risk groups)</p> <p>PTV1: CTV + 1.0–1.5cm with posterior margin of 0.5–1.0cm (both risk groups)</p> <p>PTV2: CTV + 0.0–1.0cm with posterior margin of ≤ 0.5cm (both risk groups)</p>	<p>‘1’ refers to phase 1, ‘2’ refers to phase 2</p> <p>GTV1 = pros + base SV (low risk)</p> <p>GTV1 = pros + SV (intermediate/high risk)</p> <p>CTV1 = GTV1 + 0.5cm (both risk groups)</p> <p>PTV1 = CTV1 + 0.5-1cm (both risk groups)</p> <p>GTV2 = CTV2 = PTV2 = pros only (both risk groups)</p>	<p>‘1’ refers to phase 1, ‘2’ refers to phase 2, ‘3’ refers to phase 3</p> <p>GTV = pros only (both risk groups)</p> <p>CTV1 = pros + base SV + 0.5cm (low risk)</p> <p>CTV1 = pros + SV + 0.5cm (intermediate/high risk)</p> <p>PTV1 = CTV1 + 0.5cm (both risk groups)</p> <p>CTV2 = pros + 0.5cm (low risk)</p> <p>CTV2 = pros + base SV + 0.5cm (intermediate/high risk)</p> <p>PTV2 = CTV + 0.5cm (with 0cm posteriorly, both risk groups)</p> <p>CTV3 = pros + 0cm, PTV 3 = PTV2 (both risk groups)</p>
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 retrospective review for all subjects ¹	No electronic individual plan review ²	Full prospective case reviews for the first 2 or 3 subjects at each centre ³
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

Note: This is the non-blinded version of Table 2. The blinded version can be found in a separate word document.

Table 2 The number of patients in each trial dataset, broken down by endpoint and patient and treatment related variables.

	RADAR HIGH RISK PATIENTS	RADAR INTERMEDIATE RISK PATIENTS	RT01 VALIDATION DATASET	CHHiP VALIDATION DATASET
Total subjects in dataset	205	478	388	253
PSAP events	96 (46.8%)	153 (32.0%)	176 (45.4%)	72 (28.5%)
PSAP follow-up in months (min, max, med, IQR)	(9, 121, 71, 58)	(9, 121, 84, 42)	(1, 156, 70, 74)	(2, 121, 70, 24)
OS events (deaths)	63 (30.7%)	97 (20.3%)	108 (27.8%)	41 (16.2%)
OS follow-up in months (min, max, med, IQR)	(3, 118, 77, 25)	(4, 116, 81, 23)	(5, 156, 106, 36)	(2, 140, 71, 24)
LCP/LP¹ events	47 (22.9%)	83 (17.4%)	71 (18.3%)	25 (9.9%)
LCP/LP follow-up in months (min, max, med, IQR)	(9, 123, 84, 54)	(12, 123, 84, 31)	(3, 156, 98, 61)	(2, 132, 71, 24)
Age at randomisation²	Median = 70.4 yrs	Median = 68.7 yrs	Median = 67.9 yrs	Median = 67.5 yrs
Prescribed dose	26 [66 Gy] 109 [70 Gy] 70 [74 Gy]	63 [66 Gy] 270 [70 Gy] 145 [74 Gy]	204 [64 Gy] 184 [74 Gy]	89 [57 Gy] 85 [60 Gy] 79 [74 Gy]
Disease risk group	205 [Gleason score > 7]	478 [Gleason score ≤ 7]	110 [T1b/c or T2a with (PSA + (Gleason score - 6)*10) < 15] 278 [T1b/c or T2a with (PSA + (Gleason score - 6)*10) ≥ 15 or T2b/T3a]	60 [T1b/c or T2a with PSA ≤ 10 and Gleason ≤ 6] 193 [Any of the following: Stage ≥ T2b, 10 < PSA ≤ 20, Gleason score > 6]
Cancer stage	137 [T2] 68 [T3/T4]	354 [T2] 124 [T3/T4]	235 [≤ T2a (T1b, T1c, T2a)] 153 [> T2a (T2b, T3a)]	185 [≤ T2a (T1a, T1b, T1c, T2a)] 68 [> T2a (T2b, T2c, T3a)]
Baseline PSA concentration²	Median = 15.50 ng/ml	Median = 13.90 ng/ml	Median = 13.80 ng/ml	Median = 11.70 ng/ml
Number of treatment beams	17 [3 beams] 114 [4 beams] 35 [5 beams] 19 [6 beams] 20 [≥ 7 beams]	55 [3 beams] 253 [4 beams] 56 [5 beams] 74 [6 beams] 40 [≥ 7 beams]	228 [3 beams for phase 1 of treatment] 160 [4 beams for phase 1 of treatment]	222 [≤ 4 beams] 31 [> 4 beams]
Hormone therapy duration³	93 [6 months androgen deprivation] 112 [18 months androgen deprivation]	251 [6 months androgen deprivation] 227 [18 months androgen deprivation]		

¹LCP was used as an 'estimate' of LP for RADAR, while the standard definition of LP was used for RT01 and CHHiP (see section 1 in Appendix for the definition of the LCP/LP endpoint).

²This variable was divided into two approximately equal subgroups split about the median value

³Hormone therapy duration only defined for RADAR (RT01 and CHHiP participants received 4-6 months of androgen deprivation therapy)

REFERENCES

1. 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).
2. 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).
3. 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).
4. 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): 10-year results from a randomised, phase 3, . *Lancet Oncol.* **20**, 267–281 (2019).
5. 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).
6. 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).
7. 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).
8. 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).

9. 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).

VOXEL-BASED ASSOCIATION OF DOSE AND PROGRESSION

Purpose: Reducing margins during treatment planning to decrease dose to healthy organs surrounding the prostate can risk inadequate treatment of subclinical disease. This study aimed to investigate whether lack of dose to subclinical disease is associated with increased disease progression by utilizing high-quality prostate radiotherapy clinical trial data to identify anatomically-localised regions where dose variation is associated with PSA progression (PSAP).

Methods and Materials: Planned dose distributions for 683 patients of the ‘Trial-A’s name’ were deformably registered onto a single exemplar computed tomography (CT) dataset. These were divided into high-risk and intermediate-risk sub-groups for analysis. Three independent voxel-based statistical tests, utilizing permutation testing, Cox regression modelling and LASSO feature selection, were applied to identify regions where dose variation was associated with PSAP. Results from the intermediate-risk ‘Trial-A’ sub-group were externally validated by registering dose distributions from ‘Trial-B’ (n=388) and ‘Trial-C’ (n=253) trials onto the same exemplar and repeating the tests on each of these data sets.

Results: Voxel-based Cox regression revealed regions where reduced dose was correlated with increased PSA progression. Reduced dose in regions associated with coverage at the posterior prostate, in the immediate periphery of the posterior prostate and in regions corresponding to the posterior oblique beams or posterior lateral beam boundary, was associated with increased PSAP for ‘Trial-A’ and ‘Trial-B’ patients, but not for ‘Trial-C’ patients. Reduced dose to the seminal vesicles (SV) region was also associated with increased PSAP for ‘Trial-A’ intermediate-risk patients.

Conclusions: Ensuring adequate dose coverage at the posterior prostate and immediately surrounding posterior region (including the SV), where aggressive cancer spread may be occurring, may improve tumour control. It is recommended that particular care is taken when defining margins at the prostate posterior, acknowledging the trade-off between quality of life due to rectal dose and the preferences of clinicians and patients.

INTRODUCTION

External beam radiation therapy (EBRT) is prominent in treating prostate cancer¹. The last two decades have seen increases in the precision of EBRT through new techniques such as intensity modulated radiation therapy (IMRT)² and image guided radiation therapy (IGRT)³. These enable more conformal treatments, escalated dose to the target, and decreasing toxicity in the surrounding healthy tissues^{4,5}.

However, the microscopic nature of disease in these peripheral regions is problematic. Detection is limited with current imaging technology, therefore making it subclinical, i.e., not specifically targeted in treatment. For example, extracapsular extension, in which tumour tissue has extended past the prostate's surrounding capsular layer, has been identified as subclinical disease⁶. Similarly, prostate perineural invasion, in which microscopic disease is found along or around a close-by nerve, has been shown to predict for increased disease progression, metastasis and death in dose escalated EBRT patients⁷. An incomplete identification of the distribution of disease in the prostate's immediate periphery is potentially leading to inadequate treatment.

Evidence has been sought to determine whether reduced dose in these regions is associated with measures of disease progression. Engels et al found that patients treated with implanted markers for IGRT had CTV margins reduced in the left-right direction (from 6mm to 3mm) and the anterior-posterior direction (10mm to 5mm)⁸. These patients experienced more biochemical progression, suggesting that sufficient dose in the prostate periphery is required, despite the high spatial accuracy provided by IGRT. Witte et al demonstrated that patients with biochemical or clinical progression were treated with significantly less dose (6 Gy, $p < 0.01$) in parts of the obturator region peripheral to the prostate⁹. An individual voxel, spatially registered between patients in this obturator region, was chosen for Kaplan-Meier analysis, comparing rates of post-treatment progression in patients with different doses at that voxel. Patients with less dose at this point experienced significantly more progression, with the authors concluding that patients with progression had received on average a lower dose where regional cancer spread could be expected. Chen et al similarly

found a dose-progression relationship in the obturator region¹⁰. One study from the ‘Trial-C’ trial has reported, however, no significant difference of treatment efficacy in reduced margins IGRT treatments¹¹.

No study, however, has investigated the relationship between dose and disease progression around the prostate in a comprehensive, voxel-based manner. Several voxel-based studies have investigated toxicity^{12,13,14,15,16}, but not progression. Investigating the dose-progression relationship in this manner could help locate and characterise the corresponding distribution of disease. This could provide clinicians with 3D information further enabling the optimisation of dose constraints around the prostate, informing application of appropriate margins and suitable selection of irradiation strategies. Furthermore, this analysis was conducted naively in the sense that voxels throughout the entire pelvic region (within and without the prostate region) were uniformly investigated. This provided an opportunity to also investigate broader dose-progression relationships such as how reduction in dose within the prostate itself is related to progression and/or how particular beam arrangements relate to progression.

In this study, multiple voxel-based statistical methods were employed to investigate the association between 3D planned dose and PSA progression (PSAP) in the entire pelvic anatomy. Many shortcomings have typically hindered previous voxel-based analyses^{17,10}, including misregistration of planned 3D dose distributions, false positive rates due to the large number of voxels being statistically compared, not using time-to-event data, or not controlling for intrinsic patient or treatment factors. This study utilised a combination of statistical approaches to compensate for 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, radiotherapy techniques and overall treatment approaches. ‘Validation’ was defined as applying the same voxel-based tests to datasets from two other trials to determine whether the emergent dose-progression patterns within the primary dataset were generalisable to these different external datasets. This validation also had an exploratory element, in that it identified emergent patterns in the external datasets regardless of whether they matched the patterns in the primary datasets.

METHODS AND MATERIALS

‘Trial-A’ Trial

Coordinated by the ‘Trial-A’s manager’, the ‘Trial-A’s full name’ (‘Trial-A’) phase 3 2x2 factorial trial (‘Trial-A’s code’) for locally advanced prostate cancer compared 6 months of androgen deprivation therapy (ADT) plus radiotherapy with 18 months of ADT with the same radiotherapy, with and without bisphosphonates^{18,19}. A total of 1071 recruited patients had T2 – T4 prostate cancer, undergoing dose-escalated 3D conformal EBRT with prescription doses of 66, 70 or 74 Gy in 2 Gy fractions, or 46 Gy EBRT combined with a brachytherapy boost. EBRT was delivered in up to two phases, the first delivering at least 46 Gy to a larger treatment volume (PTV1), the second delivering the remaining dose to a smaller treatment volume (PTV2) - see Table 1 for margins. Patients receiving a high dose-rate brachytherapy boost were not included in this study. Plans could be generated with any preferred combination of 3 or more conformal beams. 3D planned dose distributions with corresponding CT images including delineated CTV, rectum and bladder were collected and utilised as the primary dataset for this study. See Table 1 for information on each trial summarised for direct comparison.

‘Trial-B’ Trial

‘Trial-B’ was a phase 3, open-label, international, randomised controlled trial comparing dose-escalated conformal radiotherapy with standard-dose conformal radiotherapy^{20,21}. Accruing a total of 843 men between January 1998 and December 2001, patients had confirmed T1b – T3a prostate cancer. The patients underwent 3D conformal EBRT with either a conventional prescribed dose of 64 Gy using prescribed arrangements of either 3 or 4 beams (phase 1), or the same with an additional 4 or 6 beam boost to 74 Gy (phase 2). ADT was recommended for 6 months. Similar 3D planned dose distributions, CT and delineation data were collected and utilised as the first external validation dataset of this study.

‘Trial-C’ Trial

The ‘Trial-C’ randomised phase 3 non-inferiority trial compared conventional and hypofractionated prostate IMRT^{22,23}. A total of 3216 accrued patients, having T1b–T3a localised prostate cancer, underwent IMRT with a conventional prescribed dose of 74 Gy in 2 Gy fractions or hypofractionated courses of 60 Gy or 57 Gy in 3 Gy fractions, with optional IGRT. Patients were treated in three phases (see appendix section 7 for exact doses in each phase) with arrangements of 3 or 4 beams (anterior/lateral/posterior) or 5 beams or more if inverse planning was utilised. ADT was recommended for 6 months, but was optional for patients with low-risk disease. Similar 3D planned data was utilised as the second external validation data set for this study. Data was limited to an early cohort of ‘Trial-C’ patients with processed DICOM information available at the time of data request.

3D Data Preparation

Three CT image templates were chosen from an independent cohort of 39 prostate EBRT patients²⁴. Pairwise registrations of CT images within this cohort along with registrations between this cohort and the ‘Trial-A’ 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). See section 4 of the appendix for a CT image of each registration template.

Dose distributions were then deformed onto the templates through application of deformation vector fields obtained from the image-based registrations above. The 3D dose distributions from all phases of radiotherapy were summed together according to biologically equivalent 2 Gy per fraction dose (EQD2)²⁵, using a spatially invariant alpha/beta ratio of 3, resulting in a single distribution for each patient registered

onto each template. In the analysis, 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.

PSA Progression Endpoint Defined

The PSAP endpoint was defined as the time between the end of radiotherapy and the occurrence of the following events during post-treatment follow-up:

For 'Trial-A' and 'Trial-C' patients, a PSAP event was defined as the occurrence of biochemical progression according to the Phoenix definition (nadir + 2ng/ml)²⁶. For 'Trial-B' patients, PSAP was defined as an increase in PSA concentration to greater than the nadir by at least 50% and greater than 2 ng/ml 6 months or more after the start of radiotherapy.

Two other secondary exploratory endpoints were included in the analysis, namely overall survival (OS) and local composite progression/local progression (LCP/LP). The corresponding definitions and results pertaining to these endpoints are found in section 5 of the appendix, as they are not the central endpoints of the study. Follow-up information for all endpoints is found in Table 2.

Voxel-Based Dose Difference Permutation Test

It is recommended that Figure 1 is closely followed while reading through the following descriptions of the voxel-based tests. This test was performed according to the method outlined by Chen et al¹⁰. Following Figure 1, for each given endpoint patients were divided according to whether they experienced an endpoint 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 generating a distribution of test statistics. A threshold was derived from this distribution, used to determine the region of dose difference with statistical certainty. This method accounts for the multiple statistical 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, i.e., voxels with a mean dose difference between patients with and without an endpoint event at any desired p-value can be determined. In this study, thresholds of $p < 0.05$, $p < 0.1$, $p < 0.2$ and $p < 0.3$ were applied to thoroughly explore the dose difference (see discussion for an elaboration on this point). As shown in Figure 1, the mean dose difference map was imposed on the registration template, including the delineated CTV, bladder and rectum. If the dose difference reached statistical significance at one of the given p-value thresholds, then the voxels corresponding to this difference (the thresholded p-value map) were highlighted in green and imposed onto the dose difference map.

Uni-Voxel Cox Regression Test

This test generates a separate Cox proportional hazards model for each voxel (hence, ‘uni’-voxel), testing for association between dose in that voxel and incidence of the endpoint. Taking a given voxel, patients were divided into two groups about the median of the combined distribution of dose values, as in Figure 1. The hazard ratio (HR) of the incidence of the endpoint between the high dose value group and low dose value group was then calculated, including a corresponding p-value determining whether the HR was significantly greater than or less than 1 at the $p < 0.05$ level. When considering PSAP as the endpoint, this HR therefore compares the incidence of PSAP between each dose group, indicating the dose-progression relationship at the given voxel. Age, prescribed dose, disease risk, cancer stage, baseline PSA concentration, number of treatment beams and duration of hormone therapy were intrinsic patient or treatment factors investigated as potential control variables in each model, attempting to eliminate their confounding influence at each voxel^{27,28}. An automated selection process selected as controls only those variables that maintained a significant correlation with the endpoint (see appendix section 3). This process selected control variables at the patient level, and then incorporated these with the voxel dose variable in a Cox proportional hazards model. I.e., the same controls selected at the patient level were included for every individual voxel’s Cox model, for the given dataset. Repeating this process for every voxel produced a 3D HR map and corresponding p-value map revealing the relationship between dose and the endpoint across the pelvic anatomy. The continuous HR map was first imposed on the anatomical template. Following this,

the thresholded p-value map was imposed onto the HR map, showing (in green) voxels where $HR < 1$ or $HR > 1$ at the $p < 0.05$ level.

Multi-Voxel Cox Regression Test with LASSO Feature Selection

In contrast to the uni-voxel Cox regression test, this test combined all voxel-dose variables in the pelvic anatomy as variables in a single multivariate Cox regression model (hence, ‘multi’-voxel). The LASSO (Least Absolute Shrinkage Selection Operator²⁹) was then applied to select voxels with dose-variables that did not correlate with each other in the model, while still correlating strongly with the endpoint. 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 endpoint by minimising the partial likelihood deviance. The selected voxels were then imposed on the anatomical template, indicating whether $HR > 1$ or $HR < 1$ in each case. As with the uni-voxel Co regression test, HRs in this test compared the incidence of the endpoint (e.g. PSAP) between the high dose group and low dose group at a given voxel, with the cut-point for dose determined in the same way. The LASSO enabled selection of voxels strongly correlated with the endpoint while accounting for inter-voxel dose correlation and the multiple testing problem.

Analysis Details

Firstly, the three voxel-based tests were applied to all 683 ‘Trial-A’ patients (the “‘Trial-A’ inclusive dataset”), on all three registration templates, to test for the impact of registration. This was an internal validation to determine whether emergent dose-progression patterns on T1 would also appear on T2 and T3, thus ascertaining whether the choice of registration template impacted these patterns. The results for this component of the analysis are found in section 4 of the appendix. All subsequent components of the

analysis were conducted on the T1 template only (as the previous components conducted on T2 and T3 were sufficient for the purpose of determining the impact of registration).

Following this, the 'Trial-A' inclusive dataset was divided into intermediate-risk (Gleason score ≤ 7) and high-risk (Gleason score > 7) groups, and the voxel-based tests were repeated on these groups separately. The voxel-based tests were also applied to the 'Trial-B' and 'Trial-C' datasets, which were considered as validation datasets for the 'Trial-A' intermediate-risk dataset. They were not considered as validation datasets for the 'Trial-A' high-risk dataset as they included a smaller and incomparable proportion of high-risk patients, and were thus more comparable to the 'Trial-A' intermediate-risk dataset. These components of the analysis, with a focus on the primary PSAP endpoint, will be the emphasis of this discussion and will comprise all the results in this manuscript (with all other results in the appendix).

The voxel-based tests were also applied to a combined dataset ("COMBINED"), consisting of patients from all three trials. Results for this dataset are found in section 4 of the appendix.

All components of the above analysis were undertaken for PSAP, the primary endpoint of this study. The same were repeated for the two secondary exploratory endpoints, with results also found in section 5 of the appendix. The voxel-based dose difference permutation and uni-voxel Cox regression 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 3.8.0³⁰.

RESULTS

Trial Datasets

Table 2 shows the patient breakdown of each dataset, including patient variable and endpoint information, after patients were excluded due to loss of follow-up and missing data. The table includes this information for the intermediate and high-risk ‘Trial-A’ datasets, and the ‘Trial-B’ and ‘Trial-C’ validation datasets.

Impact of Registration

The dose-progression patterns from the ‘Trial-A’ inclusive dataset on T1 were generally reproduced on the other registration templates (T2 and T3). The patterns were distorted according to the anatomical difference between the templates, but otherwise were similar, suggesting the revealed dose-endpoint association patterns are largely independent of choice of registration template (see appendix section 5 for these results).

Results from ‘Trial-A’ Intermediate-Risk vs High-Risk Datasets

These results are found in Figure 2. For the intermediate-risk dataset, the mean dose-difference map indicates that compared to patients without PSAP, patients with PSAP had up to 2 Gy more dose on average directly superior to the prostate (see sagittal plane), and in regions corresponding to the oblique beams (see axial plane). Similarly, these patients had up to 3.5 Gy less dose on the posterior boundary of the left lateral beam. For the high-risk dataset, those with PSAP had up to 4 Gy more dose across the anterior side of the lateral beams (see axial plane), particularly manifest at the anterior side of the prostate (see axial and sagittal planes). However, the permutation test identified no significant dose difference up to the $p < 0.3$ level.

For the intermediate-risk dataset, voxel clusters (VCs) with $HR < 1$ ($p < 0.05$) were found in the oblique beam regions, particularly dominant on the posterior side (see axial plane), and in the

immediate periphery of the prostate (see axial and coronal planes), particularly on the posterior side and extending in direction of the SV (see sagittal plane). Some were also found at the left posterior lateral beam boundary. VCs with $HR > 1$ ($p < 0.05$) were found in the posterior and anterior beam regions (see axial and sagittal planes). For the high-risk dataset, VCs with $HR < 1$ ($p < 0.05$) were found across the posterior lateral beam boundary, particularly manifest on the left side (see axial plane), and in the lateral periphery of the prostate towards the posterior side of the prostate (see coronal plane). For the intermediate-risk dataset, the LASSO selected two $HR < 1$ voxel left of the prostate in the inferior direction (one seen in the axial plane, the other in the coronal plane). For the high-risk dataset, the LASSO selected no significant voxels.

In summary, the major observed associations show that reduced dose posterior to the prostate is correlated with incidence of PSAP for both the 'Trial-A' intermediate and high-risk datasets. For the intermediate-risk patients, this was most prominently manifest in posterior oblique beam regions and in the immediate periphery of the prostate, extending into the SV region. For high-risk patients, this was seen across the posterior lateral beam boundary. It is also noteworthy that results for intermediate and high-risk patients substantially differ.

Results From 'Trial-A' Intermediate-Risk Validation With 'Trial-B' and 'Trial-C' Datasets

These results are found in Figure 3. For the 'Trial-B' dataset, the mean dose-difference map indicates that compared to patients without PSAP, patients with PSAP had up to 6 Gy less dose on average across the posterior boundary of the lateral beam region (see axial plane). For the 'Trial-C' dataset, those with PSAP had up to 7 Gy more dose inferior to the prostate (see coronal and sagittal planes) and across posterior side of the lateral beams, particularly on the right side (see axial plane). However, as previously, the permutation test identified no significant dose difference up to the $p < 0.3$ level.

For the 'Trial-B' dataset, VCs with $HR < 1$ ($p < 0.05$) were found in the posterior oblique beam regions and across the posterior boundary of lateral beam region, particularly dominant on the left side (see axial plane). VCs with $HR > 1$ ($p < 0.05$) were found in the lateral beam regions (see axial and coronal planes). For the 'Trial-C' dataset, VCs with $HR < 1$ ($p < 0.05$) were found in the immediate periphery of the prostate on the left side (see axial plane) and in the anterior beam region (see axial plane). VCs with $HR > 1$ ($p < 0.05$) were found across the posterior boundary of the lateral beam region (see axial and coronal planes). No voxels were selected by the LASSO in either validation datasets.

In summary, the major observed associations show that the association between PSAP and reduced dose at the posterior boundary of the lateral beam region found in the 'Trial-A' intermediate-risk dataset was also found in the 'Trial-B' validation dataset. The association between PSAP and reduced dose and in the immediate periphery of the prostate found in the 'Trial-A' intermediate-risk dataset was also found in the 'Trial-C' dataset. The major observed association for the 'Trial-C' dataset, namely that of PSAP being associated with increased dose across the posterior lateral beam boundary, was the opposite of that found in the same region for the 'Trial-A' intermediate-risk dataset and 'Trial-B' validation dataset.

Other less prominent dose-progression association patterns are evident, but these major patterns will be the focus.

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 correlations between voxel-dose and measures of **disease-progression** across the pelvic anatomy have been identified.

Although no individual voxel-based test in this study addressed every shortcoming of voxel-based analyses, each test did address specific problems such that a consistent result across all techniques could be considered independent of these issues. The uni-voxel and multi-voxel Cox regression tests utilised post-treatment time-to-event endpoints, with the uni-voxel test controlling for **patient and treatment** factors. The LASSO regression ensured selected voxels were independent of correlation with other voxels. Incorporating all voxels in the model together accounted for the multiple comparisons problem. The dose difference permutation test similarly accounted for the multiple comparisons problem, while also being the only method of the three that excluded **noise**.

Both the intermediate and high-risk 'Trial-A' patients demonstrated an association between increased PSAP and reduced dose in the immediate periphery of the prostate, particularly on the posterior side, and even extending into the SV region for intermediate-risk patients. This association seemed to be correlated with reduced dose in posterior oblique beam regions for intermediate-risk patients, and reduced dose across the posterior boundary of the lateral beam region for high-risk patients. Minimal evidence for this association in the immediate periphery of the prostate was found for 'Trial-B' patients. However, these patients still exhibited an association between increased PSAP and reduced dose across the posterior lateral beam boundary and in the posterior oblique beam region – similar to the pattern in 'Trial-A' intermediate-risk patients. 'Trial-C' patients, however, generally did not confirm these associations. It is noteworthy that 'Trial-B' and 'Trial-C' (who's patients did not show substantial evidence for dose-progression association in the prostate periphery) included at least

the base of the SV in their definitions of GTV, while 'Trial-A' intermediate risk patients (who did show substantial evidence for dose-progression association in the prostate periphery) did not have any part of the SV included within their GTV (see Table 1). Overall, evidence has been established for a relationship between increased PSAP and reduced dose in the posterior periphery of the prostate, linked to lack of posterior oblique beams and/or posterior lateral beam coverage. This may be suggesting that reduced posterior prostate margins, and therefore reduced posterior coverage of the prostate, are associated with increased PSAP.

McNeal et al have shown that the majority of prostate cancer originates in the peripheral zone (PZ, at the prostate posterior periphery), as opposed to the transition zone (TZ, in the central anterior), with 68% arising in the PZ as opposed to 24% in the TZ³¹. Lee et al found that patients with PZ as opposed to TZ tumours had increased odds of SV invasion, extra-capsular extension, lymphovascular invasion and increased incidence of tumour recurrence³². It has also been shown that despite TZ tumours being larger at diagnosis and patients with these tumours having a higher baseline PSA concentration, PZ tumours had higher cell proliferation levels and were more associated with biomarkers related to invasive potential³³. In culmination, cancer in the PZ, at the posterior of the prostate, is more prominent and aggressively invasive than cancer originating elsewhere in the prostate. Adequate coverage at the posterior prostate and its immediate periphery, including the SV, is therefore crucial for overall tumour control. This is consistent with the prominent dose-progression pattern identified here and confirms the findings of Engels et al where reduced anterior-posterior margins were correlated with increased biochemical progression⁸. Another potential cause of this effect could be that if patient rectums were overfilled with gas or stool at planning and then emptied during treatment the prostate (and especially the SV) will move posteriorly and potentially out of the high dose volume, remembering the CT templates in this study were taken before planning. This rectal distension has been associated with increased biochemical and local progression³⁴. However, an awareness of rectal distension in the modern era of RT has enabled clinicians to address this effect³⁵. In conclusion, it is

recommended that particular care is taken when defining margins at the prostate posterior, acknowledging the trade-off between quality of life due to rectal dose and the preferences of clinicians and patients. The dose-shaping capability of IMRT and VMAT, coupled with the increased accuracy afforded by IGRT, will assist in adequate dose coverage in this posterior region.

It is peculiar that reduced dose-progression associations in the prostate posterior periphery are stronger for intermediate-risk patients than for high-risk patients. This may be attributable to the fact that the intermediate-risk dataset is larger and contains more PSAP events, and therefore more statistical power to reveal dose-progression associations. Also, the 'Trial-C' results run contrary to this general pattern. This may be attributable to somewhat more favourable prognostic features in the 'Trial-C' cohort relative to the other trials, namely a smaller proportion of patients with T3 disease and risk of SV involvement. Treatment accuracy was also expected to improve throughout the course of the 'Trial-C' trial, for example with more consistent use of cone beam CT which was not available in the 'Trial-B' era. 'Trial-A' and 'Trial-B' patients were treated more similarly in terms of RT technique, being treated in the same era. Finally, it must be noted that low-risk patients were present in the 'Trial-B' and 'Trial-C' datasets, while not in the 'Trial-A' datasets. It is recommended that future analyses exploring associating dose in the prostate periphery with progression remove low-risk patients. However, 'Trial-B'/'Trial-C' datasets both contain over 70% of patients in the intermediate/high risk category, indicating they are still adequate for validating the 'Trial-A' intermediate-risk dataset in the context of this exploratory study.

The permutation test is quite conservative. In the dose difference comparison between patients with and without an event pertaining to the given endpoint, it applies a global threshold that cannot identify local maxima of dose difference. Also, due to the large number of voxels compared, in order to adequately account for the multiple statistical testing problem this threshold can be quite high, and therefore may exclude not only local regions of significant dose difference but also global regions.

Hence only large and statistically strong global dose differences can be identified (and therefore p-value thresholds up to $p < 0.3$ were used). This could explain why, across all datasets and endpoints, only one region of statistically significant dose difference was found by this test. A test more sensitive in identifying local maxima, such as a threshold-free cluster enhancement test³⁶, may be appropriate for further voxel-based analyses.

The identified relationships are correlative and not necessarily causative, and therefore may not represent anatomically-localised physiologically caused dose-progression associations. Only the uni-voxel Cox regression accounted for intrinsic patient and treatment factors, and these represent only a sample of possible factors that could confound the associations. For example, it was expected that prescribed dose would potentially confound the relationship between voxel-dose and PSA progression. However, it was only significantly associated with PSA progression for the RT01 cohort (See Appendix section 3), and therefore only included as a control variable here. Another potential confounder could be the different prostate margins employed in each trial, which were not controlled for. Although not controlled for, the use of bisphosphonates, unique to 'Trial-A', were previously investigated and found to have no relationship with PSAP³⁷. To ensure dose-progression relationships are independent of a given patient or treatment factor, separating the cohort into this factor's subgroups prior to analysis is necessary. This, however, would reduce power, requiring a larger cohort. It is recommended that, for future studies, effort be made to collate datasets with more internal diversity across these variables, with large enough numbers of endpoint events in each variable subcategory for stratification.

This study is also limited by the assumption that planned dose is equivalent to delivered dose, which differ in practice³⁸. As the consistency between planned and delivered dose improves, or delivered dose becomes increasingly measurable, voxel-based dose analyses will become more effective in finding anatomically localised dose-endpoint relationships. Data derived from patients treated with

IGRT, for example, would ensure planned dose more closely resembles delivered dose. Additionally, voxel dose was defined as a dichotomised variable in the Cox regression. Even though this enabled a clear interpretation of the hazard ratio, it is also recommended that voxel dose be defined as a continuous variable in future analyses. Another limitation could be the accuracy of registration and the appropriateness of the choice of exemplar and anti-exemplar. A perfectly accurate registration would ensure the identified patterns are in fact occurring at the identified anatomical site. Diversity in the dose distributions across the cohort is also limiting, as the mean distributions are approximately 3 or 4 field treatments in all datasets (see appendix section 6 for mean and standard deviation distributions). Greater diversity in technique will enable more generalisable feature selection.

Biologically equivalent dose was calculated using an alpha beta ratio of 3 Gy for all voxels throughout the pelvic anatomy. Empirically determined alpha beta ratios vary greatly, being dependent on many clinical and methodological factors³⁹. For example, reported alpha beta ratios for prostate tumours vary from -0.07 Gy⁴⁰ to 18 Gy⁴¹. An appropriate future direction may be to test the sensitivity of the results to varying alpha beta ratios.

This was the first study performing voxel-based analysis of the dose-progression relationship around the prostate extending through the entire pelvic anatomy. It confirms previous work that reduced dose surrounding the posterior border of the prostate increases the risk of progression. It further reinforces the need for adequate dose coverage at the prostate posterior where aggressive cancer spread could be occurring, particularly in the SV. Translation to guiding planning might be achieved by parameterising the dose distribution to account for spatial distributions, such as through principal component analysis, functional analysis, dosiomics⁴² or a convolutional neural network approach. This will require extensively more data with more diversity.

REFERENCES

1. Greco, C. *et al.* The evolving role of external beam radiotherapy in localized prostate cancer. in *Seminars in oncology* (Elsevier, 2019).
2. Luxton, G., Hancock, S. L. & Boyer, A. L. Dosimetry and radiobiologic model comparison of IMRT and 3D conformal radiotherapy in treatment of carcinoma of the prostate. *Int. J. Radiat. Oncol. Biol. Phys.* **59**, 267–284 (2004).
3. Crevoisier, R. De *et al.* Image-guided Radiation Therapy (IGRT) in Prostate Cancer: Preliminary Results in Prostate Registration and Acute Toxicity of a Randomized Study. *Radiat. Oncol. Biol.* **75**, S99 (2009).
4. Sandler, H. M. *et al.* Reduction in Patient-reported Acute Morbidity in Prostate Cancer Patients Treated With 81-Gy Intensity-modulated Radiotherapy Using Reduced Planning Target Volume Margins and Electromagnetic Tracking: Assessing the Impact of Margin Reduction Study. *URL* **75**, 1004–1008 (2010).
5. Uang, S. H. *et al.* The Effect of Changing Technique, Dose, and PTV Margin on Therapeutic Ratio During Prostate Radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* **71**, 1057–1064 (2008).
6. Teh, B., Bastasch, M., Mai, W. & Butler, E. Predictors of extracapsular extension and its radial distance in prostate cancer: implications for prostate IMRT, brachytherapy, and surgery. *Cancer J.* **9**, 506 (2003).
7. Feng, F. Y. *et al.* Perineural Invasion Predicts Increased Recurrence, Metastasis, and Death From Prostate Cancer Following Treatment With Dose-Escalated Radiation Therapy. *Int. J. Radiat. Oncol.* **81**, 361–367 (2011).
8. Engels, B., Soete, G., Verellen, D. & Storme, G. Conformal Arc Radiotherapy for Prostate Cancer: Increased Biochemical Failure in Patients With Distended Rectum on the Planning Computed Tomogram Despite Image Guidance by Implanted Markers. *Int. J. Radiat. Oncol. Biol. Phys.* **74**, 388–391 (2009).
9. Witte, M. G. *et al.* Relating Dose Outside the Prostate With Freedom From Failure in the Dutch Trial 68 Gy vs. 78 Gy. *Int. J. Radiat. Oncol. Biol. Phys.* **77**, 131–138 (2010).
10. 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).

11. Murray, J. *et al.* A randomised assessment of image guided radiotherapy within a phase 3 trial of conventional or hypofractionated high dose intensity modulated radiotherapy for prostate cancer. *Radiother. Oncol.* **142**, 62–71 (2020).
12. 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).
13. Dréan, G. *et al.* Identification of a rectal subregion highly predictive of rectal bleeding in prostate cancer IMRT. *Radiother. Oncol.* **119**, 388–397 (2016).
14. 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).
15. Yahya, N. *et al.* Modeling Urinary Dysfunction After External Beam Radiation Therapy of the Prostate Using Bladder Dose-Surface Maps : Evidence of Spatially Variable Response of the Bladder Surface. *Radiat. Oncol. Biol.* **97**, 420–426 (2018).
16. Mylona, E. *et al.* Voxel-Based Analysis for Identification of Urethrovesical Subregions Predicting Urinary Toxicity After Prostate Cancer Radiation Therapy. *Int. J. Radiat. Oncol. Biol. Phys.* **104**, 343–354 (2019).
17. 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).
18. 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).
19. 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).
20. 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).
21. 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).

22. 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).
23. 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).
24. Kennedy, A. *et al.* Similarity clustering based atlas selection for pelvic CT image segmentation. *Med. Phys.* **46**, 2246–2250 (2019).
25. Bentzen, S. M. *et al.* Bioeffect modeling and equieffective dose concepts in radiation oncology-Terminology, quantities and units. *Radiother. Oncol.* **105**, 266–268 (2012).
26. Roach, M. *et al.* Defining biochemical failure following radiotherapy with or without hormonal therapy in men with clinically localized prostate cancer: recommendations of the RTOG-ASTRO Phoenix Consensus Conference. *Int. J. Radiat. Oncol. Biol. Phys.* **65**, 965–74 (2006).
27. VanderWeele, T. J. & Shpitser, I. On the definition of a confounder. *Ann. Stat.* **41**, 196–220 (2013).
28. 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).
29. Tibshiranit, B. R. Regression Shrinkage and Selection via the Lasso. *J. R. Stat. Soc. Ser. B* **58**, 267–288 (1996).
30. Yushkevich, P. A. *et al.* User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage* **31**, 1116–1128 (2006).
31. McNeal, J., Redwine, E., Freiha, F. & Stamey, T. Zonal distribution of prostatic adenocarcinoma: correlation with histologic pattern and direction of spread. *Am. J. Surg. Pathol.* **12**, 897–906 (1988).
32. Lee, J. J. *et al.* Biologic differences between peripheral and transition zone prostate cancer. *Prostate* **75**, 183–190 (2015).
33. Sakai, I., Harada, K., Hara, I., Eto, H. & Miyake, H. A comparison of the biological features between prostate cancers arising in the transition and peripheral zones. *BJU Int.* **96**, 528–532

- (2005).
34. de Crevoisier, R. *et al.* Increased risk of biochemical and local failure in patients with distended rectum on the planning CT for prostate cancer radiotherapy. *Int. J. Radiat. Oncol.* **62**, 965–973 (2005).
 35. Silverman, R., Johnson, K., Perry, C. & Sundar, S. Degree of rectal distension seen on prostate radiotherapy planning CT Scan is not a negative prognostic factor in the modern era of image-guided radiotherapy. *Oncology* **90**, 51–56 (2016).
 36. Smith, S. M. & Nichols, T. E. Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage* **44**, 83–98 (2009).
 37. 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): 10-year results from a randomised, phase 3, . *Lancet Oncol.* **20**, 267–281 (2019).
 38. 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).
 39. Van Leeuwen, C. M. *et al.* The alfa and beta of tumours: a review of parameters of the linear-quadratic model, derived from clinical radiotherapy studies. *Radiat. Oncol.* **13**, 96 (2018).
 40. Vogelius, I. R. & Bentzen, S. M. Meta-analysis of the alpha/beta ratio for prostate cancer in the presence of an overall time factor: bad news, good news, or no news? *Int. J. Radiat. Oncol. Biol. Phys.* **85**, 89–94 (2013).
 41. Boonstra, P. S. *et al.* Alpha/beta (α/β) ratio for prostate cancer derived from external beam radiotherapy and brachytherapy boost. *Br. J. Radiol.* **89**, 20150957 (2016).
 42. Liang, B. *et al.* Dosiomics: Extracting 3D Spatial Features From Dose Distribution to Predict Incidence of Radiation Pneumonitis. *Front. Oncol.* **9**, 1–7 (2019).

FIGURE CAPTIONS

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

Figure 2 Results from associating PSAP with voxel-dose in the 'Trial-A' intermediate-risk and high-risk datasets. Corresponding axial, coronal and sagittal slices (top to bottom) of a) mean dose difference maps, 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. 'No Voxels Selected' implies the LASSO selected no voxels of significant correlation with the endpoint. I.e., this test yielded no results. The slices chosen for display were those which coincide with the most dominant emergent dose-endpoint patterns, indicated in corresponding planes with dashed lines. Tones of red correspond to regions where increased dose is associated with incidence of PSAP ($HR > 1$), while tones of blues correspond to regions where reduced dose is associated with incidence of PSAP ($HR < 1$). The CTV is delineated in orange while the bladder and rectum are delineated in yellow. Anatomical directions left (L), right (R), superior (S), inferior (I), anterior (A), and posterior (P) are also indicated.

Figure 3 Results from associating PSAP with voxel-dose in the 'Trial-A' intermediate-risk dataset (included again for comparison), with corresponding results from the RT01 and CHHiP validation datasets. Corresponding axial, coronal and sagittal slices (top to bottom) of a) mean dose difference maps, 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. 'No Voxels Selected' implies the LASSO selected no voxels of significant correlation with the endpoint. I.e., this test yielded no results. The slices chosen for display were those which coincide with the most dominant emergent dose-endpoint patterns, indicated in corresponding planes with dashed lines. Tones of red correspond to regions where increased dose is associated with incidence of PSAP ($HR > 1$), while tones of blues correspond to regions where reduced dose is associated with incidence of PSAP ($HR < 1$). The CTV is delineated in orange while the bladder and rectum are delineated in yellow. Anatomical directions left (L), right (R), superior (S), inferior (I), anterior (A), and posterior (P) are also indicated.

VOXEL-BASED ASSOCIATION OF DOSE AND PROGRESSION

Purpose: Reducing margins during treatment planning to decrease dose to healthy organs surrounding the prostate can risk inadequate treatment of subclinical disease. This study aimed to investigate whether lack of dose to subclinical disease is associated with increased disease progression by utilizing high-quality prostate radiotherapy clinical trial data to identify anatomically-localised regions where dose variation is associated with PSA progression (PSAP).

Methods and Materials: Planned dose distributions for 683 patients of the ‘Trial-A’s name’ were deformably registered onto a single exemplar computed tomography (CT) dataset. These were divided into high-risk and intermediate-risk sub-groups for analysis. Three independent voxel-based statistical tests, utilizing permutation testing, Cox regression modelling and LASSO feature selection, were applied to identify regions where dose variation was associated with PSAP. Results from the intermediate-risk ‘Trial-A’ sub-group were externally validated by registering dose distributions from ‘Trial-B’ (n=388) and ‘Trial-C’ (n=253) trials onto the same exemplar and repeating the tests on each of these data sets.

Results: Voxel-based Cox regression revealed regions where reduced dose was correlated with increased PSA progression. Reduced dose in regions associated with coverage at the posterior prostate, in the immediate periphery of the posterior prostate and in regions corresponding to the posterior oblique beams or posterior lateral beam boundary, was associated with increased PSAP for ‘Trial-A’ and ‘Trial-B’ patients, but not for ‘Trial-C’ patients. Reduced dose to the seminal vesicles (SV) region was also associated with increased PSAP for ‘Trial-A’ intermediate-risk patients.

Conclusions: Ensuring adequate dose coverage at the posterior prostate and immediately surrounding posterior region (including the SV), where aggressive cancer spread may be occurring, may improve tumour control. It is recommended that particular care is taken when defining margins at the prostate posterior, acknowledging the trade-off between quality of life due to rectal dose and the preferences of clinicians and patients.

INTRODUCTION

External beam radiation therapy (EBRT) is prominent in treating prostate cancer¹. The last two decades have seen increases in the precision of EBRT through new techniques such as intensity modulated radiation therapy (IMRT)² and image guided radiation therapy (IGRT)³. These enable more conformal treatments, escalated dose to the target, and decreasing toxicity in the surrounding healthy tissues^{4,5}.

However, the microscopic nature of disease in these peripheral regions is problematic. Detection is limited with current imaging technology, therefore making it subclinical, i.e., not specifically targeted in treatment. For example, extracapsular extension, in which tumour tissue has extended past the prostate's surrounding capsular layer, has been identified as subclinical disease⁶. Similarly, prostate perineural invasion, in which microscopic disease is found along or around a close-by nerve, has been shown to predict for increased disease progression, metastasis and death in dose escalated EBRT patients⁷. An incomplete identification of the distribution of disease in the prostate's immediate periphery is potentially leading to inadequate treatment.

Evidence has been sought to determine whether reduced dose in these regions is associated with measures of disease progression. Engels et al found that patients treated with implanted markers for IGRT had CTV margins reduced in the left-right direction (from 6mm to 3mm) and the anterior-posterior direction (10mm to 5mm)⁸. These patients experienced more biochemical progression, suggesting that sufficient dose in the prostate periphery is required, despite the high spatial accuracy provided by IGRT. Witte et al demonstrated that patients with biochemical or clinical progression were treated with significantly less dose (6 Gy, $p < 0.01$) in parts of the obturator region peripheral to the prostate⁹. An individual voxel, spatially registered between patients in this obturator region, was chosen for Kaplan-Meier analysis, comparing rates of post-treatment progression in patients with different doses at that voxel. Patients with less dose at this point experienced significantly more progression, with the authors concluding that patients with progression had received on average a lower dose where regional cancer spread could be expected. Chen et al similarly

found a dose-progression relationship in the obturator region¹⁰. One study from the ‘Trial-C’ trial has reported, however, no significant difference of treatment efficacy in reduced margins IGRT treatments¹¹.

No study, however, has investigated the relationship between dose and disease progression around the prostate in a comprehensive, voxel-based manner. Several voxel-based studies have investigated toxicity^{12,13,14,15,16}, but not progression. Investigating the dose-progression relationship in this manner could help locate and characterise the corresponding distribution of disease. This could provide clinicians with 3D information further enabling the optimisation of dose constraints around the prostate, informing application of appropriate margins and suitable selection of irradiation strategies. Furthermore, this analysis was conducted naively in the sense that voxels throughout the entire pelvic region (within and without the prostate region) were uniformly investigated. This provided an opportunity to also investigate broader dose-progression relationships such as how reduction in dose within the prostate itself is related to progression and/or how particular beam arrangements relate to progression.

In this study, multiple voxel-based statistical methods were employed to investigate the association between 3D planned dose and PSA progression (PSAP) in the entire pelvic anatomy. Many shortcomings have typically hindered previous voxel-based analyses^{17,10}, including misregistration of planned 3D dose distributions, false positive rates due to the large number of voxels being statistically compared, not using time-to-event data, or not controlling for intrinsic patient or treatment factors. This study utilised a combination of statistical approaches to compensate for 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, radiotherapy techniques and overall treatment approaches. ‘Validation’ was defined as applying the same voxel-based tests to datasets from two other trials to determine whether the emergent dose-progression patterns within the primary dataset were generalisable to these different external datasets. This validation also had an exploratory element, in that it identified emergent patterns in the external datasets regardless of whether they matched the patterns in the primary datasets.

METHODS AND MATERIALS

‘Trial-A’ Trial

Coordinated by the ‘Trial-A’s manager’, the ‘Trial-A’s full name’ (‘Trial-A’) phase 3 2x2 factorial trial (‘Trial-A’s code’) for locally advanced prostate cancer compared 6 months of androgen deprivation therapy (ADT) plus radiotherapy with 18 months of ADT with the same radiotherapy, with and without bisphosphonates^{18,19}. A total of 1071 recruited patients had T2 – T4 prostate cancer, undergoing dose-escalated 3D conformal EBRT with prescription doses of 66, 70 or 74 Gy in 2 Gy fractions, or 46 Gy EBRT combined with a brachytherapy boost. EBRT was delivered in up to two phases, the first delivering at least 46 Gy to a larger treatment volume (PTV1), the second delivering the remaining dose to a smaller treatment volume (PTV2) - see Table 1 for margins. Patients receiving a high dose-rate brachytherapy boost were not included in this study. Plans could be generated with any preferred combination of 3 or more conformal beams. 3D planned dose distributions with corresponding CT images including delineated CTV, rectum and bladder were collected and utilised as the primary dataset for this study. See Table 1 for information on each trial summarised for direct comparison.

‘Trial-B’ Trial

‘Trial-B’ was a phase 3, open-label, international, randomised controlled trial comparing dose-escalated conformal radiotherapy with standard-dose conformal radiotherapy^{20,21}. Accruing a total of 843 men between January 1998 and December 2001, patients had confirmed T1b – T3a prostate cancer. The patients underwent 3D conformal EBRT with either a conventional prescribed dose of 64 Gy using prescribed arrangements of either 3 or 4 beams (phase 1), or the same with an additional 4 or 6 beam boost to 74 Gy (phase 2). ADT was recommended for 6 months. Similar 3D planned dose distributions, CT and delineation data were collected and utilised as the first external validation dataset of this study.

‘Trial-C’ Trial

The ‘Trial-C’ randomised phase 3 non-inferiority trial compared conventional and hypofractionated prostate IMRT^{22,23}. A total of 3216 accrued patients, having T1b–T3a localised prostate cancer, underwent IMRT with a conventional prescribed dose of 74 Gy in 2 Gy fractions or hypofractionated courses of 60 Gy or 57 Gy in 3 Gy fractions, with optional IGRT. Patients were treated in three phases (see appendix section 7 for exact doses in each phase) with arrangements of 3 or 4 beams (anterior/lateral/posterior) or 5 beams or more if inverse planning was utilised. ADT was recommended for 6 months, but was optional for patients with low-risk disease. Similar 3D planned data was utilised as the second external validation data set for this study. Data was limited to an early cohort of ‘Trial-C’ patients with processed DICOM information available at the time of data request.

3D Data Preparation

Three CT image templates were chosen from an independent cohort of 39 prostate EBRT patients²⁴. Pairwise registrations of CT images within this cohort along with registrations between this cohort and the ‘Trial-A’ 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). See section 4 of the appendix for a CT image of each registration template.

Dose distributions were then deformed onto the templates through application of deformation vector fields obtained from the image-based registrations above. The 3D dose distributions from all phases of radiotherapy were summed together according to biologically equivalent 2 Gy per fraction dose (EQD2)²⁵, using a spatially invariant alpha/beta ratio of 3, resulting in a single distribution for each patient registered

onto each template. In the analysis, 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.

PSA Progression Endpoint Defined

The PSAP endpoint was defined as the time between the end of radiotherapy and the occurrence of the following events during post-treatment follow-up:

For ‘Trial-A’ and ‘Trial-C’ patients, a PSAP event was defined as the occurrence of biochemical progression according to the Phoenix definition (nadir + 2ng/ml)²⁶. For ‘Trial-B’ patients, PSAP was defined as an increase in PSA concentration to greater than the nadir by at least 50% and greater than 2 ng/ml 6 months or more after the start of radiotherapy.

Two other secondary exploratory endpoints were included in the analysis, namely overall survival (OS) and local composite progression/local progression (LCP/LP). The corresponding definitions and results pertaining to these endpoints are found in section 5 of the appendix, as they are not the central endpoints of the study. Follow-up information for all endpoints is found in Table 2.

Voxel-Based Dose Difference Permutation Test

It is recommended that Figure 1 is closely followed while reading through the following descriptions of the voxel-based tests. This test was performed according to the method outlined by Chen et al¹⁰. Following Figure 1, for each given endpoint patients were divided according to whether they experienced an endpoint 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 generating a distribution of test statistics. A threshold was derived from this distribution, used to determine the region of dose difference with statistical certainty. This method accounts for the multiple statistical 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, i.e., voxels with a mean dose difference between patients with and without an endpoint event at any desired p-value can be determined. In this study, thresholds of $p < 0.05$, $p < 0.1$, $p < 0.2$ and $p < 0.3$ were applied to thoroughly explore the dose difference (see discussion for an elaboration on this point). As shown in Figure 1, the mean dose difference map was imposed on the registration template, including the delineated CTV, bladder and rectum. If the dose difference reached statistical significance at one of the given p-value thresholds, then the voxels corresponding to this difference (the thresholded p-value map) were highlighted in green and imposed onto the dose difference map.

Uni-Voxel Cox Regression Test

This test generates a separate Cox proportional hazards model for each voxel (hence, ‘uni’-voxel), testing for association between dose in that voxel and incidence of the endpoint. Taking a given voxel, patients were divided into two groups about the median of the combined distribution of dose values, as in Figure 1. The hazard ratio (HR) of the incidence of the endpoint between the high dose value group and low dose value group was then calculated, including a corresponding p-value determining whether the HR was significantly greater than or less than 1 at the $p < 0.05$ level. When considering PSAP as the endpoint, this HR therefore compares the incidence of PSAP between each dose group, indicating the dose-progression relationship at the given voxel. Age, prescribed dose, disease risk, cancer stage, baseline PSA concentration, number of treatment beams and duration of hormone therapy were intrinsic patient or treatment factors investigated as potential control variables in each model, attempting to eliminate their confounding influence at each voxel^{27,28}. An automated selection process selected as controls only those variables that maintained a significant correlation with the endpoint (see appendix section 3). This process selected control variables at the patient level, and then incorporated these with the voxel dose variable in a Cox proportional hazards model. I.e., the same controls selected at the patient level were included for every individual voxel’s Cox model, for the given dataset. Repeating this process for every voxel produced a 3D HR map and corresponding p-value map revealing the relationship between dose and the endpoint across the pelvic anatomy. The continuous HR map was first imposed on the anatomical template. Following this,

the thresholded p-value map was imposed onto the HR map, showing (in green) voxels where $HR < 1$ or $HR > 1$ at the $p < 0.05$ level.

Multi-Voxel Cox Regression Test with LASSO Feature Selection

In contrast to the uni-voxel Cox regression test, this test combined all voxel-dose variables in the pelvic anatomy as variables in a single multivariate Cox regression model (hence, ‘multi’-voxel). The LASSO (Least Absolute Shrinkage Selection Operator²⁹) was then applied to select voxels with dose-variables that did not correlate with each other in the model, while still correlating strongly with the endpoint. 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 endpoint by minimising the partial likelihood deviance. The selected voxels were then imposed on the anatomical template, indicating whether $HR > 1$ or $HR < 1$ in each case. As with the uni-voxel Co regression test, HRs in this test compared the incidence of the endpoint (e.g. PSAP) between the high dose group and low dose group at a given voxel, with the cut-point for dose determined in the same way. The LASSO enabled selection of voxels strongly correlated with the endpoint while accounting for inter-voxel dose correlation and the multiple testing problem.

Analysis Details

Firstly, the three voxel-based tests were applied to all 683 ‘Trial-A’ patients (the “‘Trial-A’ inclusive dataset”), on all three registration templates, to test for the impact of registration. This was an internal validation to determine whether emergent dose-progression patterns on T1 would also appear on T2 and T3, thus ascertaining whether the choice of registration template impacted these patterns. The results for this component of the analysis are found in section 4 of the appendix. All subsequent components of the

analysis were conducted on the T1 template only (as the previous components conducted on T2 and T3 were sufficient for the purpose of determining the impact of registration).

Following this, the ‘Trial-A’ inclusive dataset was divided into intermediate-risk (Gleason score ≤ 7) and high-risk (Gleason score > 7) groups, and the voxel-based tests were repeated on these groups separately. The voxel-based tests were also applied to the ‘Trial-B’ and ‘Trial-C’ datasets, which were considered as validation datasets for the ‘Trial-A’ intermediate-risk dataset. They were not considered as validation datasets for the ‘Trial-A’ high-risk dataset as they included a smaller and incomparable proportion of high-risk patients, and were thus more comparable to the ‘Trial-A’ intermediate-risk dataset. These components of the analysis, with a focus on the primary PSAP endpoint, will be the emphasis of this discussion and will comprise all the results in this manuscript (with all other results in the appendix).

The voxel-based tests were also applied to a combined dataset (“COMBINED”), consisting of patients from all three trials. Results for this dataset are found in section 4 of the appendix.

All components of the above analysis were undertaken for PSAP, the primary endpoint of this study. The same were repeated for the two secondary exploratory endpoints, with results also found in section 5 of the appendix. The voxel-based dose difference permutation and uni-voxel Cox regression 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 3.8.0³⁰.

RESULTS

Trial Datasets

Table 2 shows the patient breakdown of each dataset, including patient variable and endpoint information, after patients were excluded due to loss of follow-up and missing data. The table includes this information for the intermediate and high-risk ‘Trial-A’ datasets, and the ‘Trial-B’ and ‘Trial-C’ validation datasets.

Impact of Registration

The dose-progression patterns from the ‘Trial-A’ inclusive dataset on T1 were generally reproduced on the other registration templates (T2 and T3). The patterns were distorted according to the anatomical difference between the templates, but otherwise were similar, suggesting the revealed dose-endpoint association patterns are largely independent of choice of registration template (see appendix section 5 for these results).

Results from ‘Trial-A’ Intermediate-Risk vs High-Risk Datasets

These results are found in Figure 2. For the intermediate-risk dataset, the mean dose-difference map indicates that compared to patients without PSAP, patients with PSAP had up to 2 Gy more dose on average directly superior to the prostate (see sagittal plane), and in regions corresponding to the oblique beams (see axial plane). Similarly, these patients had up to 3.5 Gy less dose on the posterior boundary of the left lateral beam. For the high-risk dataset, those with PSAP had up to 4 Gy more dose across the anterior side of the lateral beams (see axial plane), particularly manifest at the anterior side of the prostate (see axial and sagittal planes). However, the permutation test identified no significant dose difference up to the $p < 0.3$ level.

For the intermediate-risk dataset, voxel clusters (VCs) with $HR < 1$ ($p < 0.05$) were found in the oblique beam regions, particularly dominant on the posterior side (see axial plane), and in the

immediate periphery of the prostate (see axial and coronal planes), particularly on the posterior side and extending in direction of the SV (see sagittal plane). Some were also found at the left posterior lateral beam boundary. VCs with $HR > 1$ ($p < 0.05$) were found in the posterior and anterior beam regions (see axial and sagittal planes). For the high-risk dataset, VCs with $HR < 1$ ($p < 0.05$) were found across the posterior lateral beam boundary, particularly manifest on the left side (see axial plane), and in the lateral periphery of the prostate towards the posterior side of the prostate (see coronal plane). For the intermediate-risk dataset, the LASSO selected two $HR < 1$ voxel left of the prostate in the inferior direction (one seen in the axial plane, the other in the coronal plane). For the high-risk dataset, the LASSO selected no significant voxels.

In summary, the major observed associations show that reduced dose posterior to the prostate is correlated with incidence of PSAP for both the ‘Trial-A’ intermediate and high-risk datasets. For the intermediate-risk patients, this was most prominently manifest in posterior oblique beam regions and in the immediate periphery of the prostate, extending into the SV region. For high-risk patients, this was seen across the posterior lateral beam boundary. It is also noteworthy that results for intermediate and high-risk patients substantially differ.

Results From ‘Trial-A’ Intermediate-Risk Validation With ‘Trial-B’ and ‘Trial-C’ Datasets

These results are found in Figure 3. For the ‘Trial-B’ dataset, the mean dose-difference map indicates that compared to patients without PSAP, patients with PSAP had up to 6 Gy less dose on average across the posterior boundary of the lateral beam region (see axial plane). For the ‘Trial-C’ dataset, those with PSAP had up to 7 Gy more dose inferior to the prostate (see coronal and sagittal planes) and across posterior side of the lateral beams, particularly on the right side (see axial plane). However, as previously, the permutation test identified no significant dose difference up to the $p < 0.3$ level.

For the ‘Trial-B’ dataset, VCs with $HR < 1$ ($p < 0.05$) were found in the posterior oblique beam regions and across the posterior boundary of lateral beam region, particularly dominant on the left side (see axial plane). VCs with $HR > 1$ ($p < 0.05$) were found in the lateral beam regions (see axial and coronal planes). For the ‘Trial-C’ dataset, VCs with $HR < 1$ ($p < 0.05$) were found in the immediate periphery of the prostate on the left side (see axial plane) and in the anterior beam region (see axial plane). VCs with $HR > 1$ ($p < 0.05$) were found across the posterior boundary of the lateral beam region (see axial and coronal planes). No voxels were selected by the LASSO in either validation datasets.

In summary, the major observed associations show that the association between PSAP and reduced dose at the posterior boundary of the lateral beam region found in the ‘Trial-A’ intermediate-risk dataset was also found in the ‘Trial-B’ validation dataset. The association between PSAP and reduced dose and in the immediate periphery of the prostate found in the ‘Trial-A’ intermediate-risk dataset was also found in the ‘Trial-C’ dataset. The major observed association for the ‘Trial-C’ dataset, namely that of PSAP being associated with increased dose across the posterior lateral beam boundary, was the opposite of that found in the same region for the ‘Trial-A’ intermediate-risk dataset and ‘Trial-B’ validation dataset.

Other less prominent dose-progression association patterns are evident, but these major patterns will be the focus.

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 correlations between voxel-dose and measures of disease-progression across the pelvic anatomy have been identified.

Although no individual voxel-based test in this study addressed every shortcoming of voxel-based analyses, each test did address specific problems such that a consistent result across all techniques could be considered independent of these issues. The uni-voxel and multi-voxel Cox regression tests utilised post-treatment time-to-event endpoints, with the uni-voxel test controlling for patient and treatment factors. The LASSO regression ensured selected voxels were independent of correlation with other voxels. Incorporating all voxels in the model together accounted for the multiple comparisons problem. The dose difference permutation test similarly accounted for the multiple comparisons problem, while also being the only method of the three that excluded noise.

Both the intermediate and high-risk 'Trial-A' patients demonstrated an association between increased PSAP and reduced dose in the immediate periphery of the prostate, particularly on the posterior side, and even extending into the SV region for intermediate-risk patients. This association seemed to be correlated with reduced dose in posterior oblique beam regions for intermediate-risk patients, and reduced dose across the posterior boundary of the lateral beam region for high-risk patients. Minimal evidence for this association in the immediate periphery of the prostate was found for 'Trial-B' patients. However, these patients still exhibited an association between increased PSAP and reduced dose across the posterior lateral beam boundary and in the posterior oblique beam region – similar to the pattern in 'Trial-A' intermediate-risk patients. 'Trial-C' patients, however, generally did not confirm these associations. It is noteworthy that 'Trial-B' and 'Trial-C' (who's patients did not show substantial evidence for dose-progression association in the prostate periphery) included at least

the base of the SV in their definitions of GTV, while 'Trial-A' intermediate risk patients (who did show substantial evidence for dose-progression association in the prostate periphery) did not have any part of the SV included within their GTV (see Table 1). Overall, evidence has been established for a relationship between increased PSAP and reduced dose in the posterior periphery of the prostate, linked to lack of posterior oblique beams and/or posterior lateral beam coverage. This may be suggesting that reduced posterior prostate margins, and therefore reduced posterior coverage of the prostate, are associated with increased PSAP.

McNeal et al have shown that the majority of prostate cancer originates in the peripheral zone (PZ, at the prostate posterior periphery), as opposed to the transition zone (TZ, in the central anterior), with 68% arising in the PZ as opposed to 24% in the TZ³¹. Lee et al found that patients with PZ as opposed to TZ tumours had increased odds of SV invasion, extra-capsular extension, lymphovascular invasion and increased incidence of tumour recurrence³². It has also been shown that despite TZ tumours being larger at diagnosis and patients with these tumours having a higher baseline PSA concentration, PZ tumours had higher cell proliferation levels and were more associated with biomarkers related to invasive potential³³. In culmination, cancer in the PZ, at the posterior of the prostate, is more prominent and aggressively invasive than cancer originating elsewhere in the prostate. Adequate coverage at the posterior prostate and its immediate periphery, including the SV, is therefore crucial for overall tumour control. This is consistent with the prominent dose-progression pattern identified here and confirms the findings of Engels et al where reduced anterior-posterior margins were correlated with increased biochemical progression⁸. Another potential cause of this effect could be that if patient rectums were overfilled with gas or stool at planning and then emptied during treatment the prostate (and especially the SV) will move posteriorly and potentially out of the high dose volume, remembering the CT templates in this study were taken before planning. This rectal distension has been associated with increased biochemical and local progression³⁴. However, an awareness of rectal distension in the modern era of RT has enabled clinicians to address this effect³⁵. In conclusion, it is

recommended that particular care is taken when defining margins at the prostate posterior, acknowledging the trade-off between quality of life due to rectal dose and the preferences of clinicians and patients. The dose-shaping capability of IMRT and VMAT, coupled with the increased accuracy afforded by IGRT, will assist in adequate dose coverage in this posterior region.

It is peculiar that reduced dose-progression associations in the prostate posterior periphery are stronger for intermediate-risk patients than for high-risk patients. This may be attributable to the fact that the intermediate-risk dataset is larger and contains more PSAP events, and therefore more statistical power to reveal dose-progression associations. Also, the 'Trial-C' results run contrary to this general pattern. This may be attributable to somewhat more favourable prognostic features in the 'Trial-C' cohort relative to the other trials, namely a smaller proportion of patients with T3 disease and risk of SV involvement. Treatment accuracy was also expected to improve throughout the course of the 'Trial-C' trial, for example with more consistent use of cone beam CT which was not available in the 'Trial-B' era. 'Trial-A' and 'Trial-B' patients were treated more similarly in terms of RT technique, being treated in the same era. Finally, it must be noted that low-risk patients were present in the 'Trial-B' and 'Trial-C' datasets, while not in the 'Trial-A' datasets. It is recommended that future analyses exploring associating dose in the prostate periphery with progression remove low-risk patients. However, 'Trial-B'/'Trial-C' datasets both contain over 70% of patients in the intermediate/high risk category, indicating they are still adequate for validating the 'Trial-A' intermediate-risk dataset in the context of this exploratory study.

The permutation test is quite conservative. In the dose difference comparison between patients with and without an event pertaining to the given endpoint, it applies a global threshold that cannot identify local maxima of dose difference. Also, due to the large number of voxels compared, in order to adequately account for the multiple statistical testing problem this threshold can be quite high, and therefore may exclude not only local regions of significant dose difference but also global regions.

Hence only large and statistically strong global dose differences can be identified (and therefore p-value thresholds up to $p < 0.3$ were used). This could explain why, across all datasets and endpoints, only one region of statistically significant dose difference was found by this test. A test more sensitive in identifying local maxima, such as a threshold-free cluster enhancement test³⁶, may be appropriate for further voxel-based analyses.

The identified relationships are correlative and not necessarily causative, and therefore may not represent anatomically-localised physiologically caused dose-progression associations. Only the uni-voxel Cox regression accounted for intrinsic patient and treatment factors, and these represent only a sample of possible factors that could confound the associations. For example, it was expected that prescribed dose would potentially confound the relationship between voxel-dose and PSA progression. However, it was only significantly associated with PSA progression for the RT01 cohort (See Appendix section 3), and therefore only included as a control variable here. Another potential confounder could be the different prostate margins employed in each trial, which were not controlled for. Although not controlled for, the use of bisphosphonates, unique to 'Trial-A', were previously investigated and found to have no relationship with PSAP³⁷. To ensure dose-progression relationships are independent of a given patient or treatment factor, separating the cohort into this factor's subgroups prior to analysis is necessary. This, however, would reduce power, requiring a larger cohort. It is recommended that, for future studies, effort be made to collate datasets with more internal diversity across these variables, with large enough numbers of endpoint events in each variable subcategory for stratification.

This study is also limited by the assumption that planned dose is equivalent to delivered dose, which differ in practice³⁸. As the consistency between planned and delivered dose improves, or delivered dose becomes increasingly measurable, voxel-based dose analyses will become more effective in finding anatomically localised dose-endpoint relationships. Data derived from patients treated with

IGRT, for example, would ensure planned dose more closely resembles delivered dose. Additionally, voxel dose was defined as a dichotomised variable in the Cox regression. Even though this enabled a clear interpretation of the hazard ratio, it is also recommended that voxel dose be defined as a continuous variable in future analyses. Another limitation could be the accuracy of registration and the appropriateness of the choice of exemplar and anti-exemplar. A perfectly accurate registration would ensure the identified patterns are in fact occurring at the identified anatomical site. Diversity in the dose distributions across the cohort is also limiting, as the mean distributions are approximately 3 or 4 field treatments in all datasets (see appendix section 6 for mean and standard deviation distributions). Greater diversity in technique will enable more generalisable feature selection.

Biologically equivalent dose was calculated using an alpha beta ratio of 3 Gy for all voxels throughout the pelvic anatomy. Empirically determined alpha beta ratios vary greatly, being dependent on many clinical and methodological factors³⁹. For example, reported alpha beta ratios for prostate tumours vary from -0.07 Gy⁴⁰ to 18 Gy⁴¹. An appropriate future direction may be to test the sensitivity of the results to varying alpha beta ratios.

This was the first study performing voxel-based analysis of the dose-progression relationship around the prostate extending through the entire pelvic anatomy. It confirms previous work that reduced dose surrounding the posterior border of the prostate increases the risk of progression. It further reinforces the need for adequate dose coverage at the prostate posterior where aggressive cancer spread could be occurring, particularly in the SV. Translation to guiding planning might be achieved by parameterising the dose distribution to account for spatial distributions, such as through principal component analysis, functional analysis, dosiomics⁴² or a convolutional neural network approach. This will require extensively more data with more diversity.

REFERENCES

1. Greco, C. *et al.* The evolving role of external beam radiotherapy in localized prostate cancer. in *Seminars in oncology* (Elsevier, 2019).
2. Luxton, G., Hancock, S. L. & Boyer, A. L. Dosimetry and radiobiologic model comparison of IMRT and 3D conformal radiotherapy in treatment of carcinoma of the prostate. *Int. J. Radiat. Oncol. Biol. Phys.* **59**, 267–284 (2004).
3. Crevoisier, R. De *et al.* Image-guided Radiation Therapy (IGRT) in Prostate Cancer: Preliminary Results in Prostate Registration and Acute Toxicity of a Randomized Study. *Radiat. Oncol. Biol.* **75**, S99 (2009).
4. Sandler, H. M. *et al.* Reduction in Patient-reported Acute Morbidity in Prostate Cancer Patients Treated With 81-Gy Intensity-modulated Radiotherapy Using Reduced Planning Target Volume Margins and Electromagnetic Tracking: Assessing the Impact of Margin Reduction Study. *URL* **75**, 1004–1008 (2010).
5. Uang, S. H. *et al.* The Effect of Changing Technique, Dose, and PTV Margin on Therapeutic Ratio During Prostate Radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* **71**, 1057–1064 (2008).
6. Teh, B., Bastasch, M., Mai, W. & Butler, E. Predictors of extracapsular extension and its radial distance in prostate cancer: implications for prostate IMRT, brachytherapy, and surgery. *Cancer J.* **9**, 506 (2003).
7. Feng, F. Y. *et al.* Perineural Invasion Predicts Increased Recurrence, Metastasis, and Death From Prostate Cancer Following Treatment With Dose-Escalated Radiation Therapy. *Int. J. Radiat. Oncol.* **81**, 361–367 (2011).
8. Engels, B., Soete, G., Verellen, D. & Storme, G. Conformal Arc Radiotherapy for Prostate Cancer: Increased Biochemical Failure in Patients With Distended Rectum on the Planning Computed Tomogram Despite Image Guidance by Implanted Markers. *Int. J. Radiat. Oncol. Biol. Phys.* **74**, 388–391 (2009).
9. Witte, M. G. *et al.* Relating Dose Outside the Prostate With Freedom From Failure in the Dutch Trial 68 Gy vs. 78 Gy. *Int. J. Radiat. Oncol. Biol. Phys.* **77**, 131–138 (2010).
10. 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).

11. Murray, J. *et al.* A randomised assessment of image guided radiotherapy within a phase 3 trial of conventional or hypofractionated high dose intensity modulated radiotherapy for prostate cancer. *Radiother. Oncol.* **142**, 62–71 (2020).
12. 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).
13. Dréan, G. *et al.* Identification of a rectal subregion highly predictive of rectal bleeding in prostate cancer IMRT. *Radiother. Oncol.* **119**, 388–397 (2016).
14. 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).
15. Yahya, N. *et al.* Modeling Urinary Dysfunction After External Beam Radiation Therapy of the Prostate Using Bladder Dose-Surface Maps : Evidence of Spatially Variable Response of the Bladder Surface. *Radiat. Oncol. Biol.* **97**, 420–426 (2018).
16. Mylona, E. *et al.* Voxel-Based Analysis for Identification of Urethrovesical Subregions Predicting Urinary Toxicity After Prostate Cancer Radiation Therapy. *Int. J. Radiat. Oncol. Biol. Phys.* **104**, 343–354 (2019).
17. 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).
18. 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).
19. 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).
20. 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).
21. 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).

22. 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).
23. 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).
24. Kennedy, A. *et al.* Similarity clustering based atlas selection for pelvic CT image segmentation. *Med. Phys.* **46**, 2246–2250 (2019).
25. Bentzen, S. M. *et al.* Bioeffect modeling and equieffective dose concepts in radiation oncology-Terminology, quantities and units. *Radiother. Oncol.* **105**, 266–268 (2012).
26. Roach, M. *et al.* Defining biochemical failure following radiotherapy with or without hormonal therapy in men with clinically localized prostate cancer: recommendations of the RTOG-ASTRO Phoenix Consensus Conference. *Int. J. Radiat. Oncol. Biol. Phys.* **65**, 965–74 (2006).
27. VanderWeele, T. J. & Shpitser, I. On the definition of a confounder. *Ann. Stat.* **41**, 196–220 (2013).
28. 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).
29. Tibshiranit, B. R. Regression Shrinkage and Selection via the Lasso. *J. R. Stat. Soc. Ser. B* **58**, 267–288 (1996).
30. Yushkevich, P. A. *et al.* User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage* **31**, 1116–1128 (2006).
31. McNeal, J., Redwine, E., Freiha, F. & Stamey, T. Zonal distribution of prostatic adenocarcinoma: correlation with histologic pattern and direction of spread. *Am. J. Surg. Pathol.* **12**, 897–906 (1988).
32. Lee, J. J. *et al.* Biologic differences between peripheral and transition zone prostate cancer. *Prostate* **75**, 183–190 (2015).
33. Sakai, I., Harada, K., Hara, I., Eto, H. & Miyake, H. A comparison of the biological features between prostate cancers arising in the transition and peripheral zones. *BJU Int.* **96**, 528–532

- (2005).
34. de Crevoisier, R. *et al.* Increased risk of biochemical and local failure in patients with distended rectum on the planning CT for prostate cancer radiotherapy. *Int. J. Radiat. Oncol.* **62**, 965–973 (2005).
 35. Silverman, R., Johnson, K., Perry, C. & Sundar, S. Degree of rectal distension seen on prostate radiotherapy planning CT Scan is not a negative prognostic factor in the modern era of image-guided radiotherapy. *Oncology* **90**, 51–56 (2016).
 36. Smith, S. M. & Nichols, T. E. Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage* **44**, 83–98 (2009).
 37. 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): 10-year results from a randomised, phase 3, . *Lancet Oncol.* **20**, 267–281 (2019).
 38. 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).
 39. Van Leeuwen, C. M. *et al.* The alfa and beta of tumours: a review of parameters of the linear-quadratic model, derived from clinical radiotherapy studies. *Radiat. Oncol.* **13**, 96 (2018).
 40. Vogelius, I. R. & Bentzen, S. M. Meta-analysis of the alpha/beta ratio for prostate cancer in the presence of an overall time factor: bad news, good news, or no news? *Int. J. Radiat. Oncol. Biol. Phys.* **85**, 89–94 (2013).
 41. Boonstra, P. S. *et al.* Alpha/beta (α/β) ratio for prostate cancer derived from external beam radiotherapy and brachytherapy boost. *Br. J. Radiol.* **89**, 20150957 (2016).
 42. Liang, B. *et al.* Dosiomics: Extracting 3D Spatial Features From Dose Distribution to Predict Incidence of Radiation Pneumonitis. *Front. Oncol.* **9**, 1–7 (2019).

FIGURE CAPTIONS

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

Figure 2 Results from associating PSAP with voxel-dose in the 'Trial-A' intermediate-risk and high-risk datasets. Corresponding axial, coronal and sagittal slices (top to bottom) of a) mean dose difference maps, 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. 'No Voxels Selected' implies the LASSO selected no voxels of significant correlation with the endpoint. I.e., this test yielded no results. The slices chosen for display were those which coincide with the most dominant emergent dose-endpoint patterns, indicated in corresponding planes with dashed lines. Tones of red correspond to regions where increased dose is associated with incidence of PSAP ($HR > 1$), while tones of blues correspond to regions where reduced dose is associated with incidence of PSAP ($HR < 1$). The CTV is delineated in orange while the bladder and rectum are delineated in yellow. Anatomical directions left (L), right (R), superior (S), inferior (I), anterior (A), and posterior (P) are also indicated.

Figure 3 Results from associating PSAP with voxel-dose in the 'Trial-A' intermediate-risk dataset (included again for comparison), with corresponding results from the RT01 and CHHiP validation datasets. Corresponding axial, coronal and sagittal slices (top to bottom) of a) mean dose difference maps, 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. 'No Voxels Selected' implies the LASSO selected no voxels of significant correlation with the endpoint. I.e., this test yielded no results. The slices chosen for display were those which coincide with the most dominant emergent dose-endpoint patterns, indicated in corresponding planes with dashed lines. Tones of red correspond to regions where increased dose is associated with incidence of PSAP ($HR > 1$), while tones of blues correspond to regions where reduced dose is associated with incidence of PSAP ($HR < 1$). The CTV is delineated in orange while the bladder and rectum are delineated in yellow. Anatomical directions left (L), right (R), superior (S), inferior (I), anterior (A), and posterior (P) are also indicated.

Note: The non-blinded version of this table is found on the Title Page. Table 1 is also only found on the Title Page and not in a separate document as it was almost impossible to present a separate blinded version of this table as it contains information concerning the trials all the way through.

Table 2 The number of patients in each trial dataset, broken down by endpoint and patient and treatment related variables.

	‘Trial-A’ HIGH RISK PATIENTS	‘Trial-A’ INTERMEDIATE RISK PATIENTS	‘Trial-B’ VALIDATION DATASET	‘Trial-C’ VALIDATION DATASET
Total subjects in dataset	205	478	388	253
PSAP events	96 (46.8%)	153 (32.0%)	176 (45.4%)	72 (28.5%)
PSAP follow-up in months (min, max, med, IQR)	(9, 121, 71, 58)	(9, 121, 84, 42)	(1, 156, 70, 74)	(2, 121, 70, 24)
OS events (deaths)	63 (30.7%)	97 (20.3%)	108 (27.8%)	41 (16.2%)
OS follow-up in months (min, max, med, IQR)	(3, 118, 77, 25)	(4, 116, 81, 23)	(5, 156, 106, 36)	(2, 140, 71, 24)
LCP/LP¹ events	47 (22.9%)	83 (17.4%)	71 (18.3%)	25 (9.9%)
LCP/LP follow-up in months (min, max, med, IQR)	(9, 123, 84, 54)	(12, 123, 84, 31)	(3, 156, 98, 61)	(2, 132, 71, 24)
Age at randomisation²	Median = 70.4 yrs	Median = 68.7 yrs	Median = 67.9 yrs	Median = 67.5 yrs
Prescribed dose	26 [66 Gy] 109 [70 Gy] 70 [74 Gy]	63 [66 Gy] 270 [70 Gy] 145 [74 Gy]	204 [64 Gy] 184 [74 Gy]	89 [57 Gy] 85 [60 Gy] 79 [74 Gy]
Disease risk group	205 [Gleason score > 7]	478 [Gleason score ≤ 7]	110 [T1b/c or T2a with (PSA + (Gleason score - 6)*10) < 15] 278 [T1b/c or T2a with (PSA + (Gleason score - 6)*10) ≥ 15 or T2b/T3a]	60 [T1b/c or T2a with PSA ≤ 10 and Gleason ≤ 6] 193 [Any of the following: Stage ≥ T2b, 10 < PSA ≤ 20, Gleason score > 6]
Cancer stage	137 [T2] 68 [T3/T4]	354 [T2] 124 [T3/T4]	235 [≤ T2a (T1b, T1c, T2a)] 153 [> T2a (T2b, T3a)]	185 [≤ T2a (T1a, T1b, T1c, T2a)] 68 [> T2a (T2b, T2c, T3a)]
Baseline PSA concentration²	Median = 15.50 ng/ml	Median = 13.90 ng/ml	Median = 13.80 ng/ml	Median = 11.70 ng/ml
Number of treatment beams	17 [3 beams] 114 [4 beams] 35 [5 beams] 19 [6 beams] 20 [≥ 7 beams]	55 [3 beams] 253 [4 beams] 56 [5 beams] 74 [6 beams] 40 [≥ 7 beams]	228 [3 beams for phase 1 of treatment] 160 [4 beams for phase 1 of treatment]	222 [≤ 4 beams] 31 [> 4 beams]
Hormone therapy duration³	93 [6 months androgen deprivation] 112 [18 months androgen deprivation]	251 [6 months androgen deprivation] 227 [18 months androgen deprivation]		

¹LCP was used as an ‘estimate’ of LP for ‘Trial-A’, while the standard definition of LP was used for ‘Trial-B’ and ‘Trial-C’ (see section 1 in Appendix for the definition of the LCP/LP endpoint).

²This variable was divided into two approximately equal subgroups split about the median value

³Hormone therapy duration only defined for ‘Trial-A’ (‘Trial-B’ and ‘Trial-C’ participants received 4-6 months of androgen deprivation therapy)

‘Trial-A’ (T1) Intermediate Risk Patients [Gleason score ≤ 7]‘Trial-A’ (T1) High Risk Patients [Gleason score > 7]

Total: 478, Events: 153 (32.0%)

Total: 205, Events: 96 (46.8%)

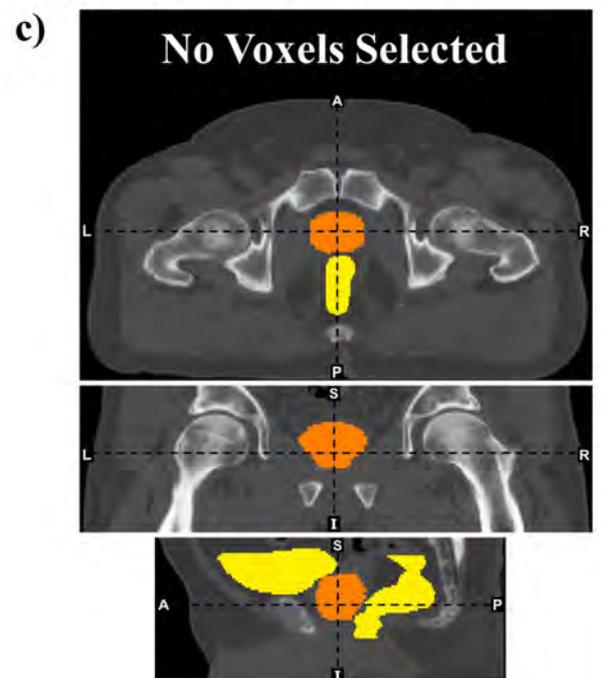
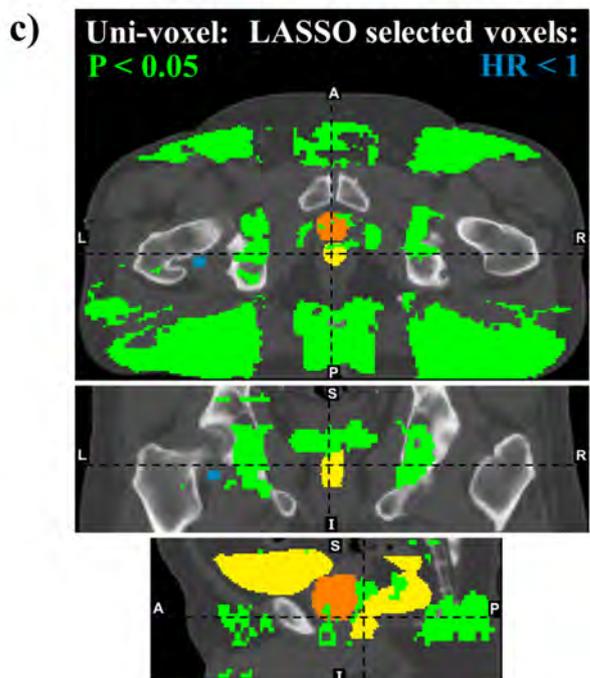
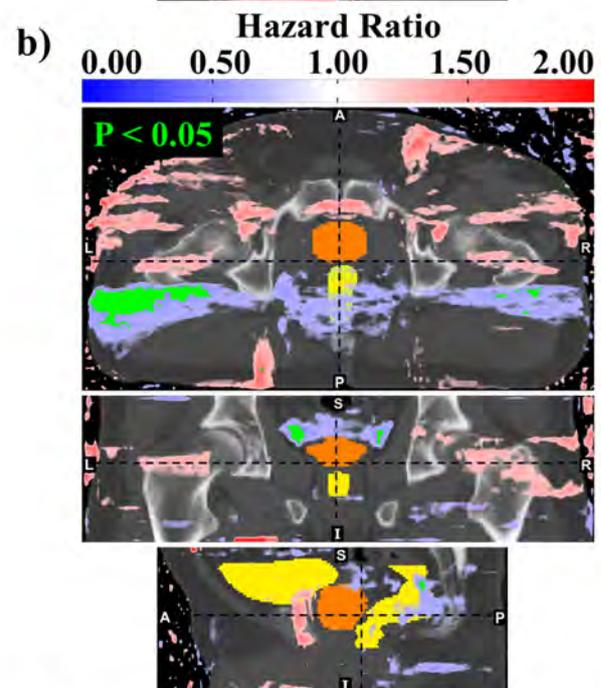
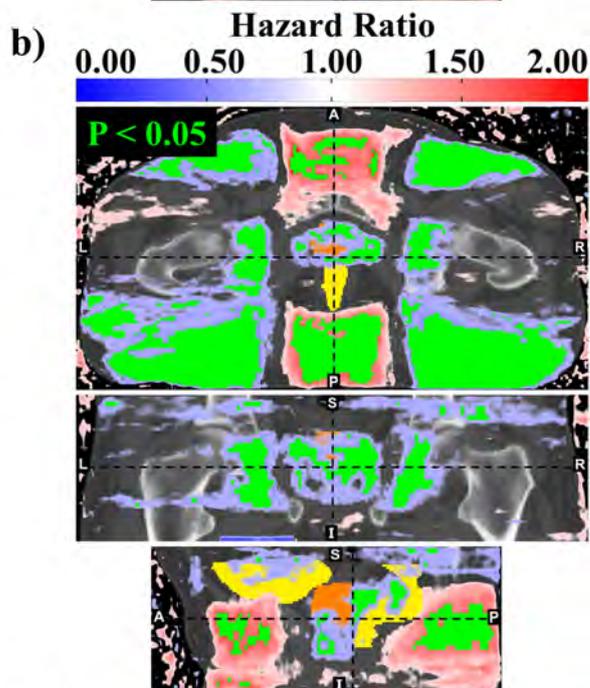
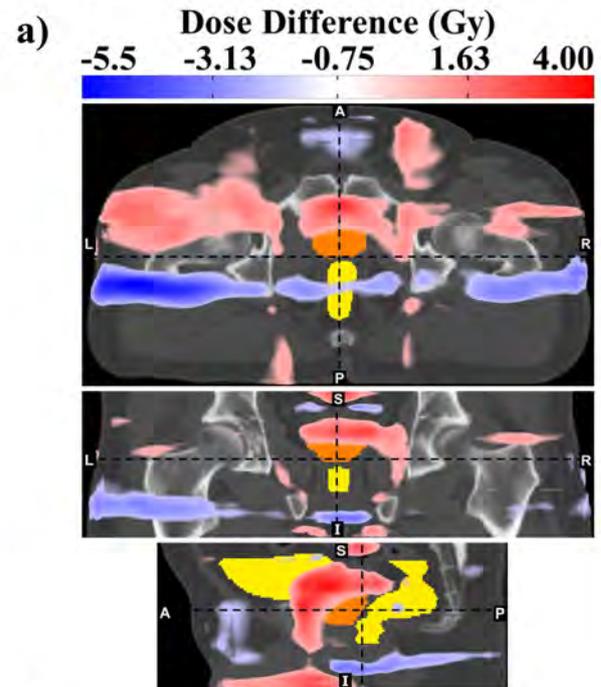
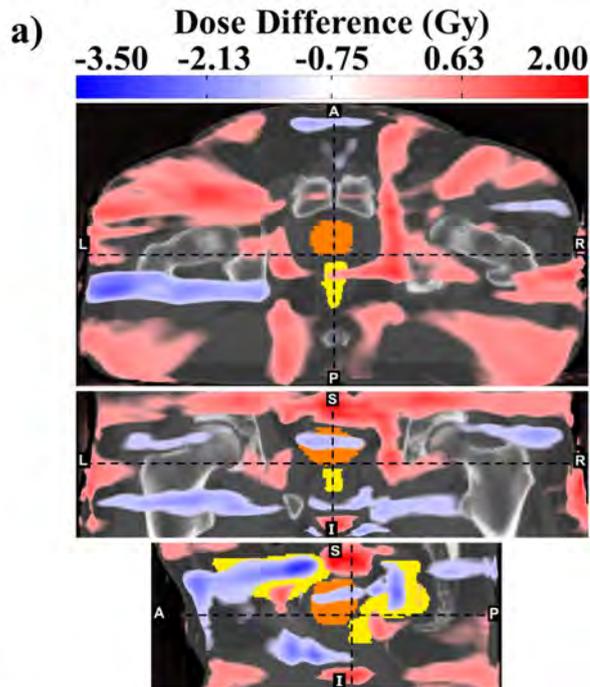
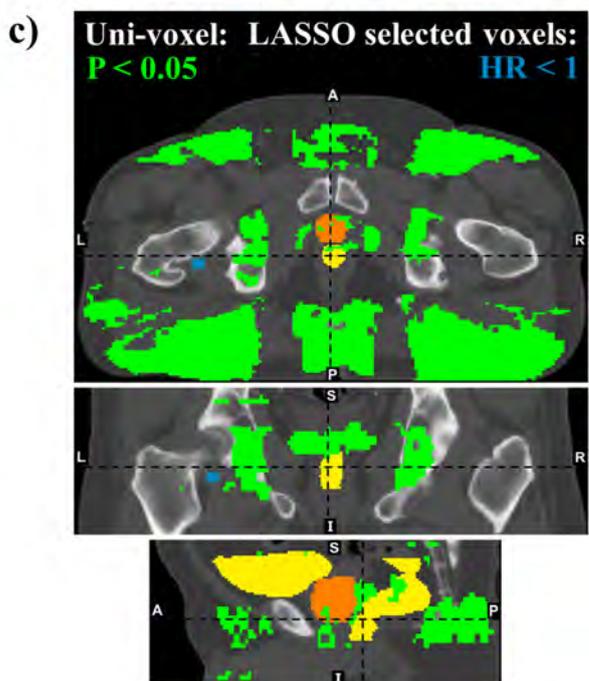
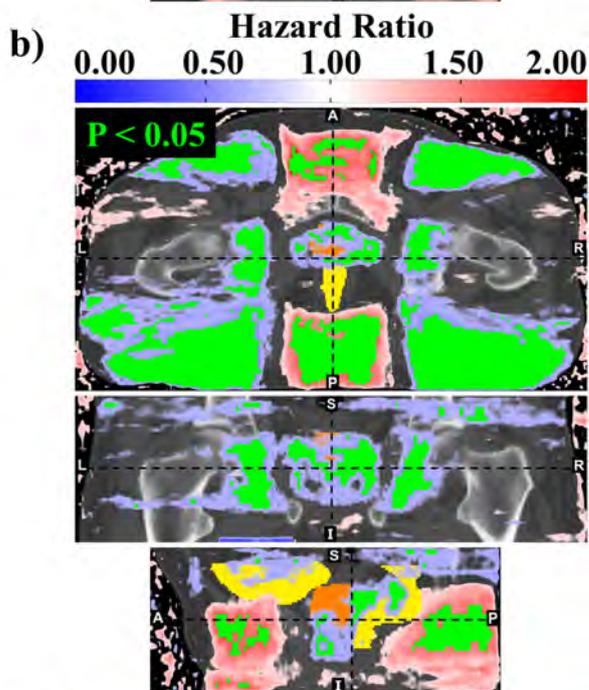
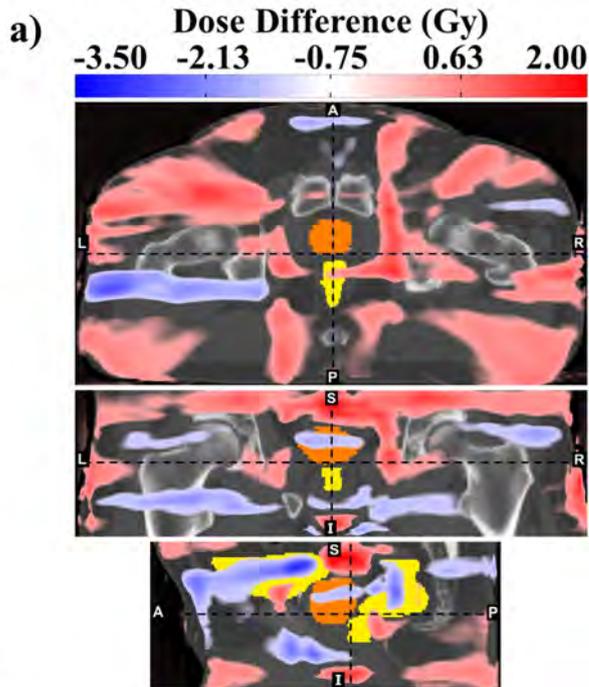


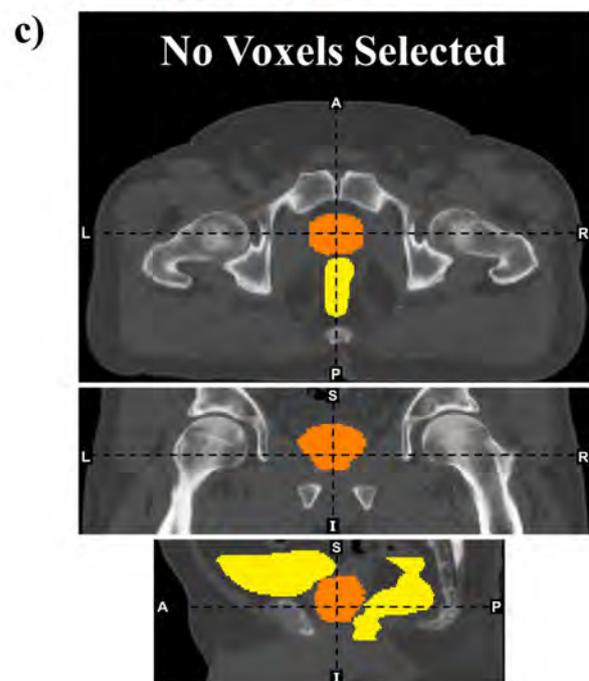
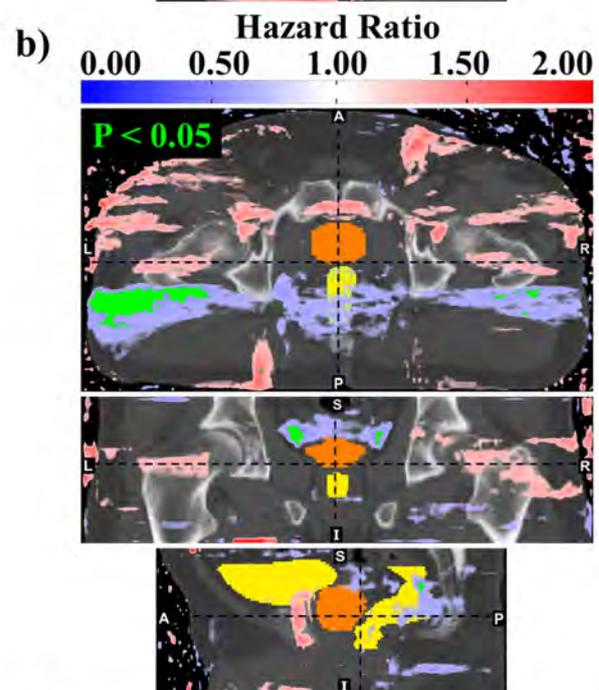
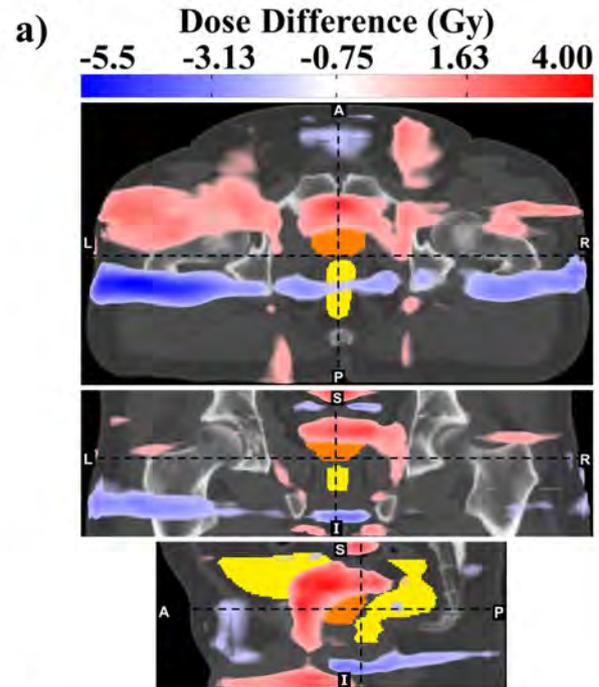
Figure 2

[Click here to access/download;Figure;Figure 2.pdf](#)
RADAR (T1) Intermediate Risk Patients [Gleason score ≤ 7]

Total: 478, Events: 153 (32.0%)


RADAR (T1) High Risk Patients [Gleason score > 7]

Total: 205, Events: 96 (46.8%)



'Trial-A' (T1)
Intermediate Risk Patients [Gleason score ≤ 7]
Total: 478, Events: 153 (32.0%)

'Trial-B' (T1)
Validation Dataset
Total: 388, Events: 176 (45.5%)

'Trial-C' (T1)
Validation Dataset
Total: 253, Events: 72 (28.5%)

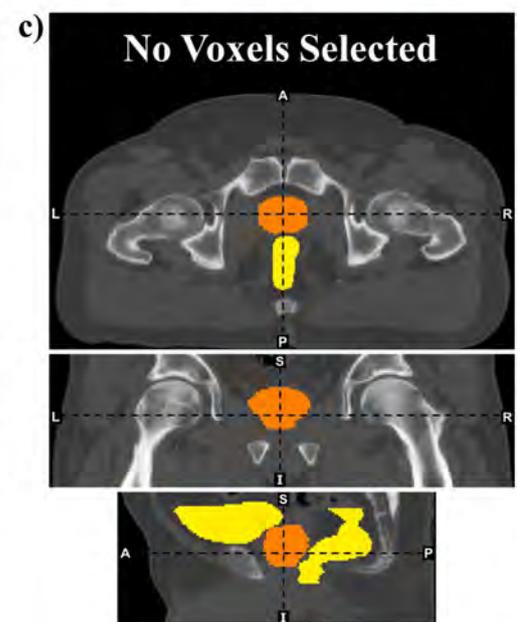
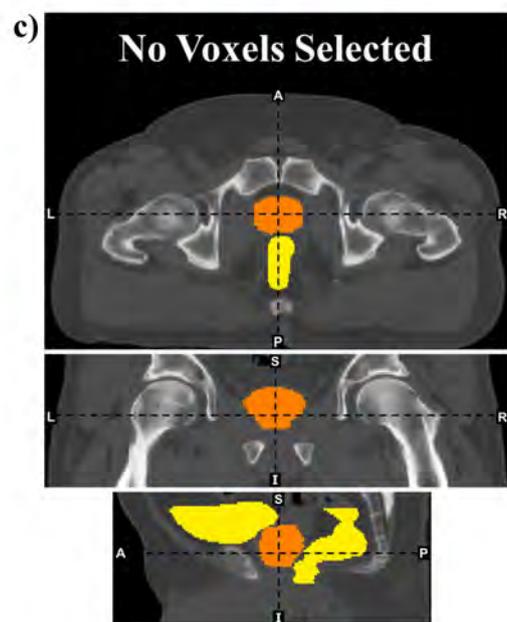
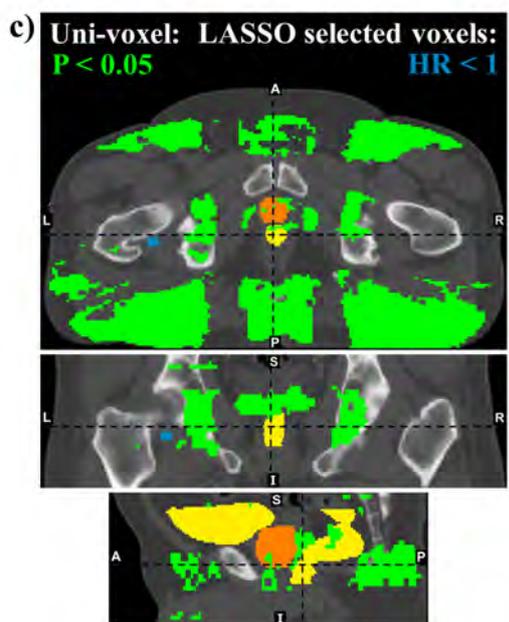
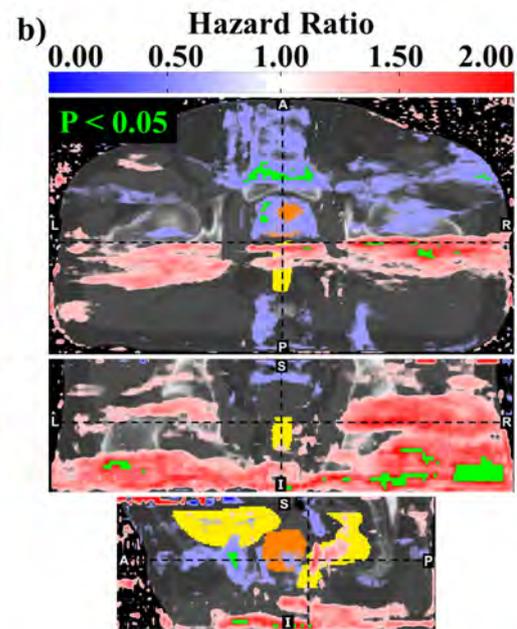
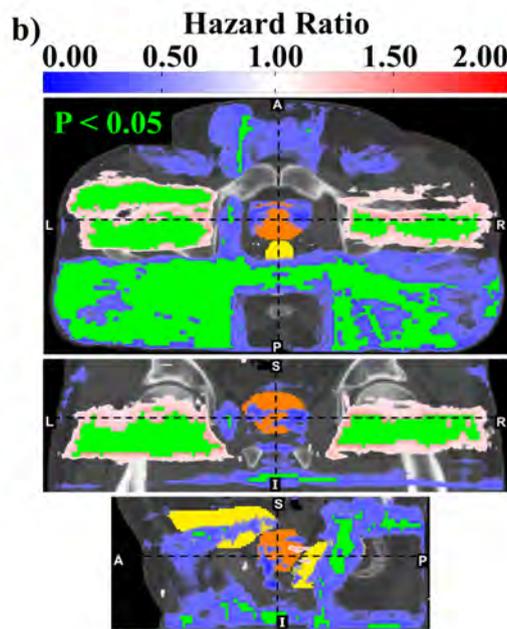
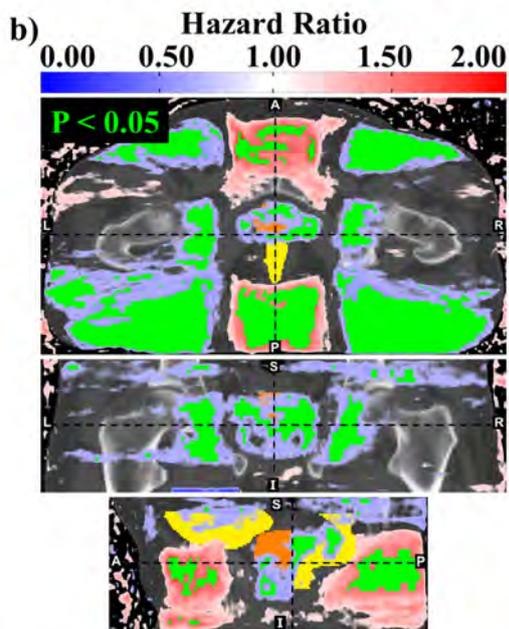
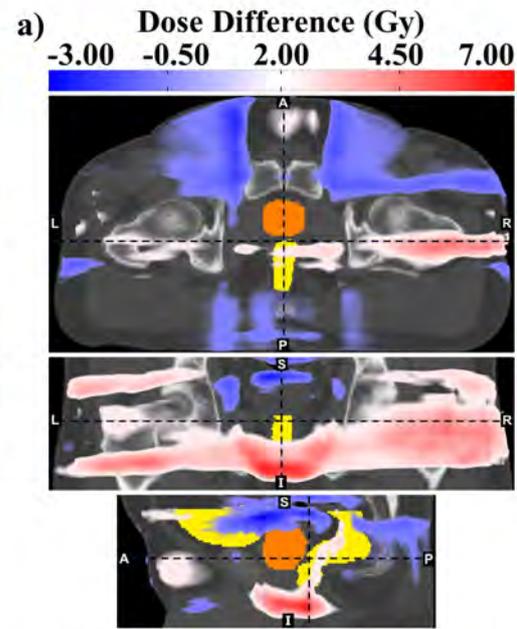
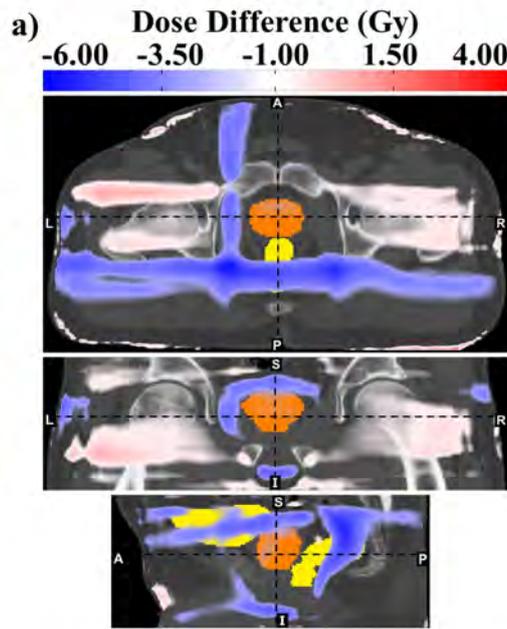
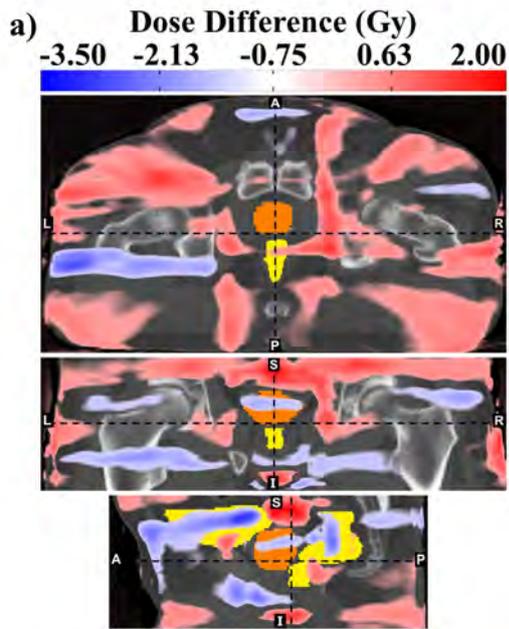


Figure 3

[Click here to access/download;Figure;Figure 3.pdf](#)

RADAR (T1)

Intermediate Risk Patients [Gleason score ≤ 7]

Total: 478, Events: 153 (32.0%)

RT01 (T1)

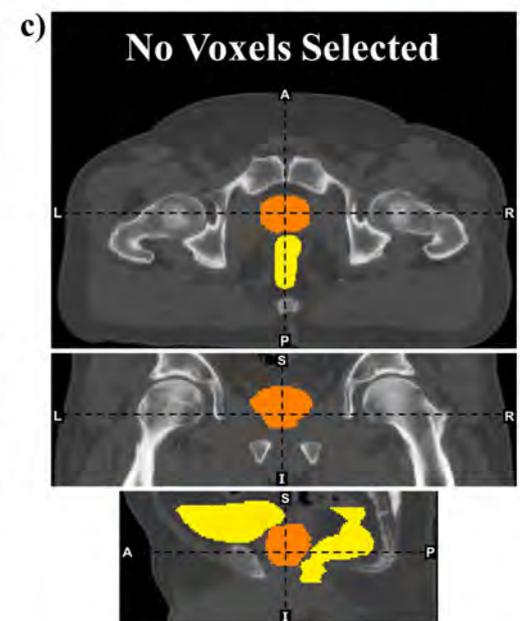
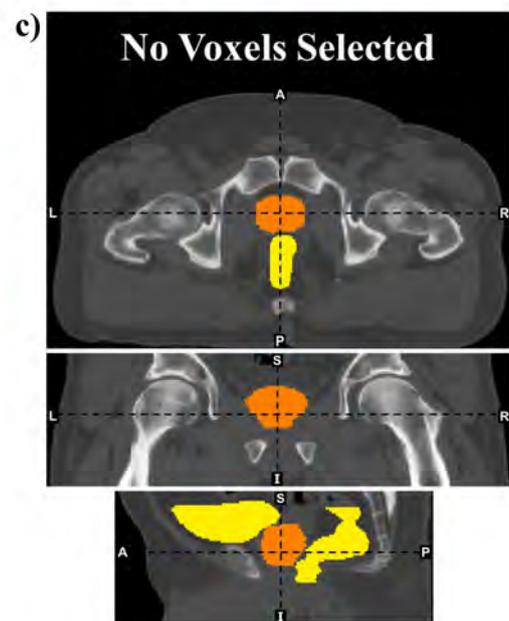
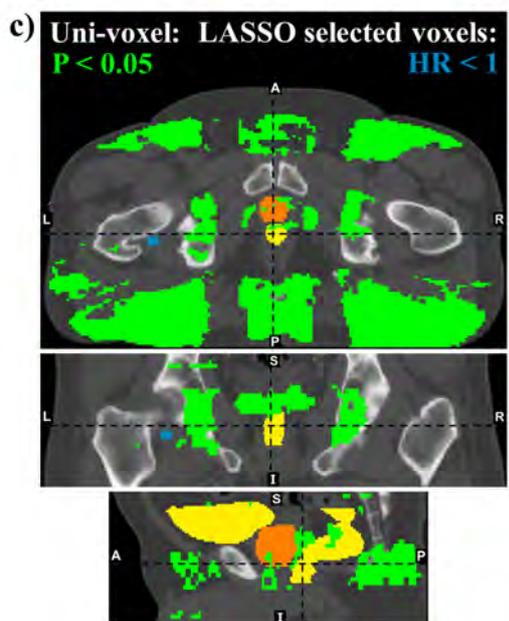
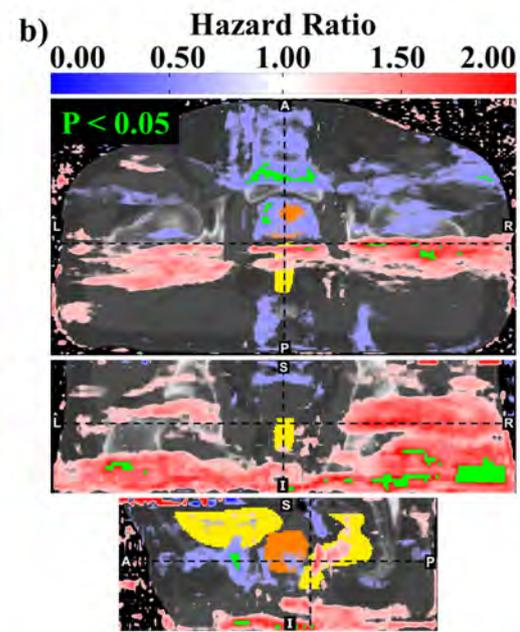
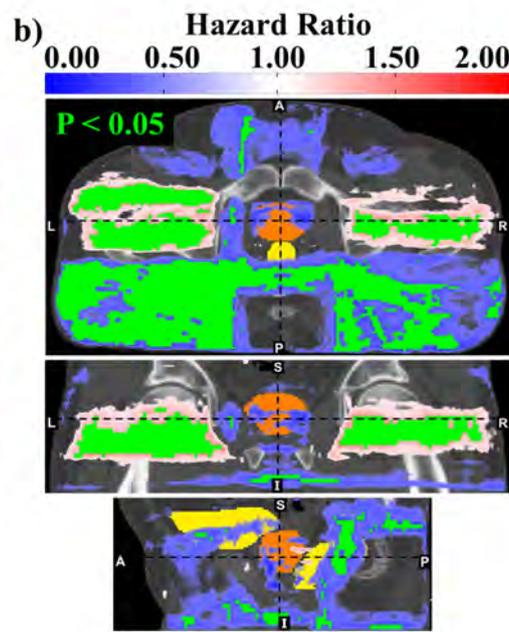
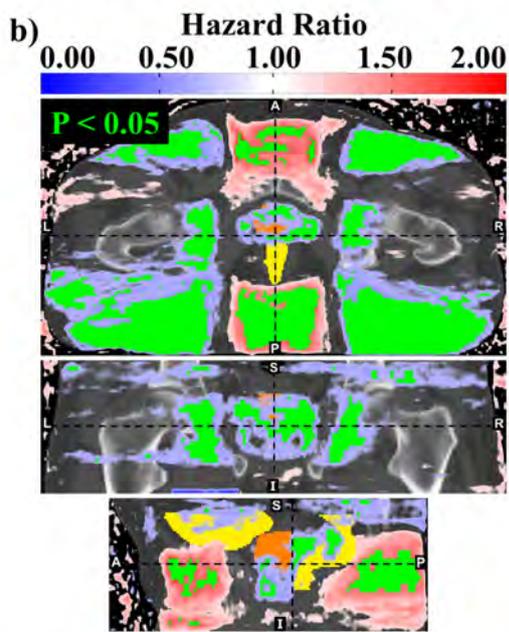
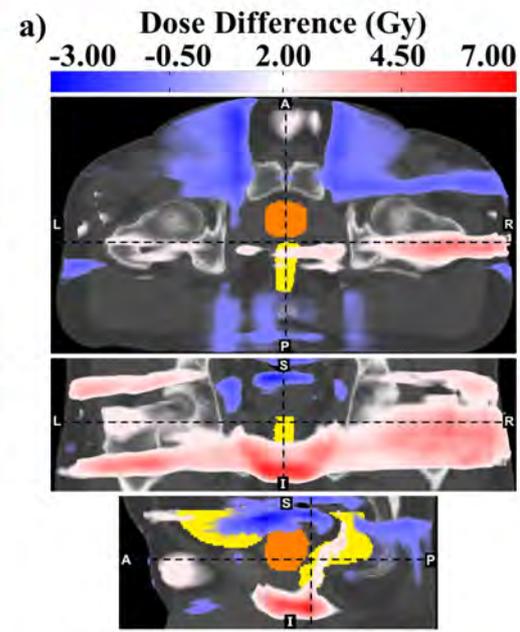
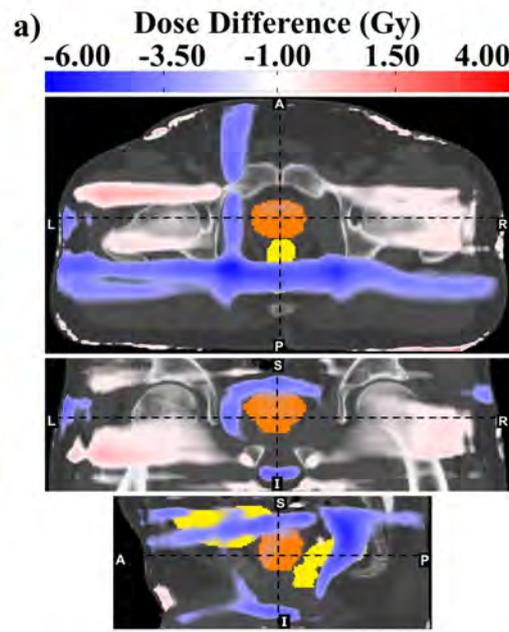
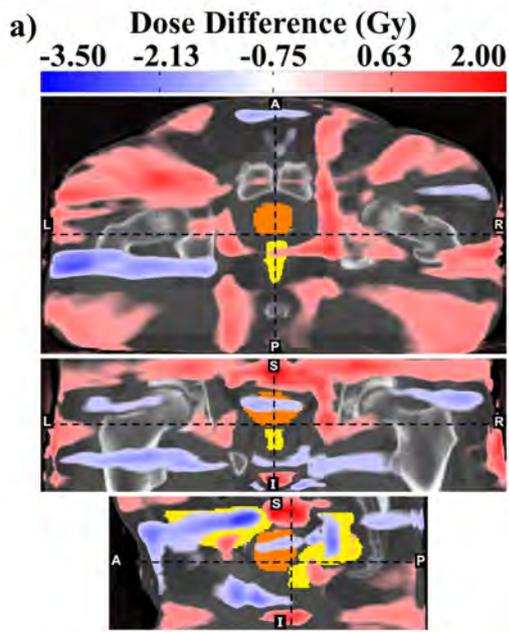
Validation Dataset

Total: 388, Events: 176 (45.5%)

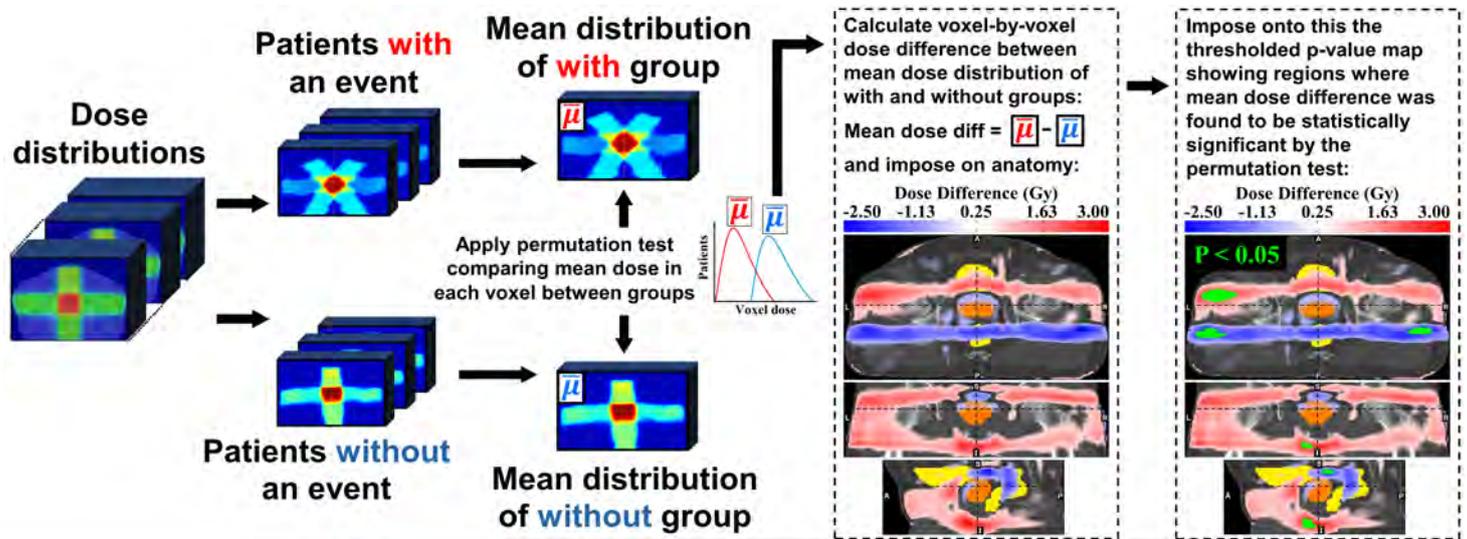
CHHiP (T1)

Validation Dataset

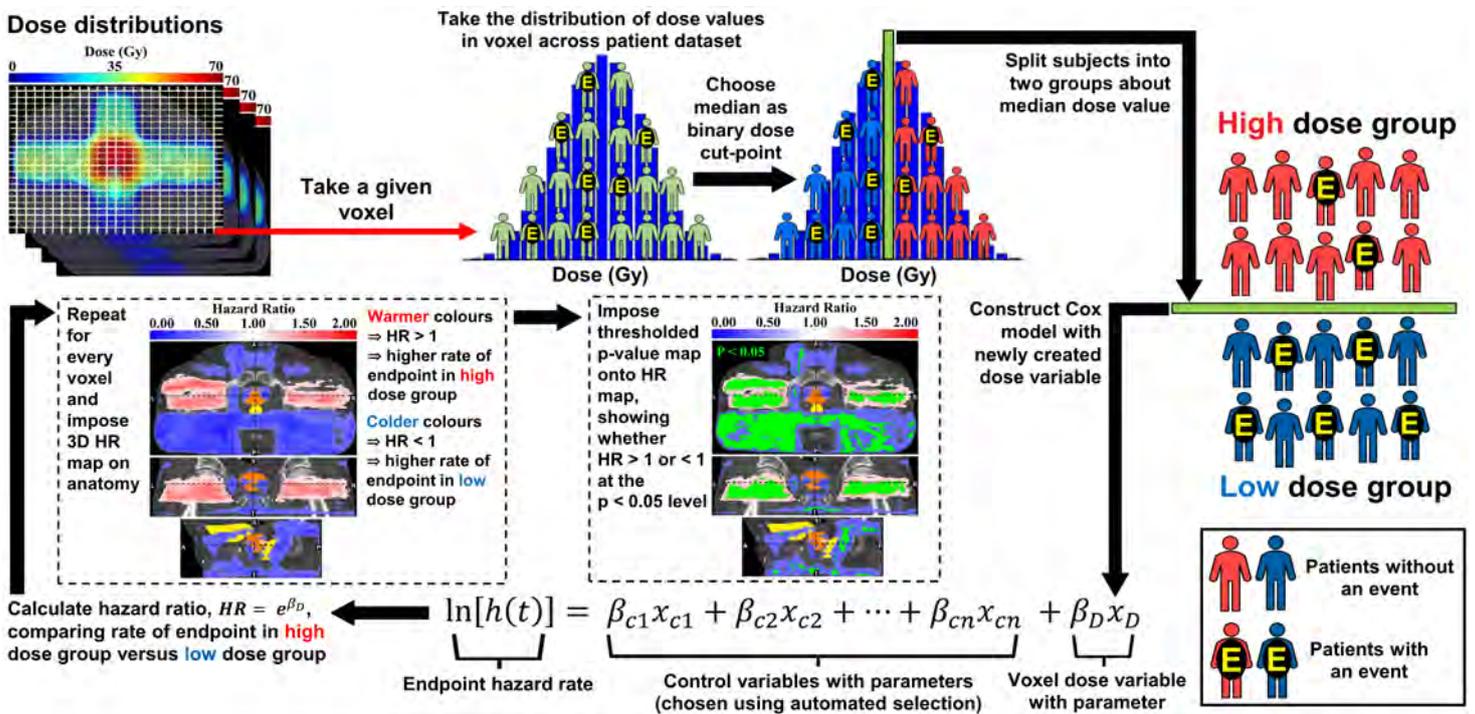
Total: 253, Events: 72 (28.5%)



a) Voxel-Based Dose Difference Permutation Test



b) Uni-Voxel Cox Regression Test



c) Multi-Voxel Cox Regression Test with LASSO Feature Selection

