INTERNATIONAL JOURNAL OF
RADIATION ONCOLOGY • BIOLOGY • PHYSICS
www.redjournal.org

PHYSICS CONTRIBUTION

Genitourinary α/β Ratios in the CHHiP Trial the Fraction Size Sensitivity of Late Genitourinary Toxicity: Analysis of Alpha/Beta (α/β) Ratios in the CHHiP Trial

Douglas H. Brand, PhD,[*,†,‡] Sarah C. Brüningk, PhD,[§,‖] Anna Wilkins, PhD,[*,†] Olivia Naismith, MSc,[¶] Annie Gao, MSc,[*,†] Isabel Syndikus, MD,[#] David P. Dearnaley, FRCR,[*,†] Nicholas van As, FRCR,[*,†] Emma Hall, PhD,[**] Sarah Gulliford, PhD,[‡,††] and Alison C. Tree, MD(Res)[*,†] On behalf of the CHHiP Trial Management Group

[*]*Division of Radiotherapy and Imaging, The Institute of Cancer Research, London, United Kingdom;* [†]*Urology Unit, Royal Marsden NHS Foundation Trust, London, United Kingdom;* [‡]*Department of Medical Physics and Biomedical Engineering, University College London, London, United Kingdom;* [§]*Department of Biosystems Science and Engineering, ETH Zurich, Basel, Switzerland;* [‖]*Swiss Institute for Bioinformatics (SIB), Lausanne, Switzerland;* [¶]*Radiotherapy Trials QA Group (RTTQA), Royal Marsden NHS Foundation Trust, London, United Kingdom;* [#]*Radiotherapy Department, Clatterbridge Cancer Centre, Liverpool, United Kingdom;* [**]*Clinical Trials and Statistics Unit, The Institute of Cancer Research, London, United Kingdom; and* [††]*Department of Radiotherapy Physics,*

Corresponding author: Douglas Brand; E-mail: douglas.brand@icr.ac.uk
Sarah Gulliford and Alison C. Tree contributed equally to this work.

CHHiP was funded by Cancer Research UK (C8262/A7253, C1491/A9895, C1491/A15955, C1491/A25351, SP2312/021), the Department of Health and the National Institute for Health Research (NIHR) Cancer Research Network. We acknowledge support from the NIHR Biomedical Research Centre at the Royal Marsden NHS Foundation Trust and The Institute of Cancer Research, London, UK. We acknowledge Cancer Research UK Programme Grants during the time of research C33589/A19727, C1491/A25351, and C33589/A28284. D.H.B. was funded by a Cancer Research UK Clinical Research Training Fellowship. Both D.H.B. and S.G. are now supported by the Cancer Research UK RadNet City of London Centre Award (C7893/A28990). The work of S.B. is supported by the Botnar Research Centre for Child Health Postdoctoral Excellence Programme under PEP-2021-1008. A.C.T. acknowledges support from CRUK grants C33589/A28284 and C7224/A28724. A.W. acknowledges support from Cancer Research UK RadNet Institute of Cancer Research/Royal Marsden Hospital.

Disclosures: D.H.B. reports Cancer Research UK fellowship funding which supported this work. D.P.D. declares personal fees for advisory board participation and consultancy from Janssen; personal fees from the Institute of Cancer Research (from the Rewards to Discoverer's Scheme for abiraterone) and has a patent (EP1933709B1, localization and stabilization device) issued. A.T. reports grants, personal fees and other from Elekta, grants and personal fees from Accuray, grants from Varian, other from Janssen, outside the submitted work; E.H. reports grants from Varian Medical Systems Inc, Astra Zeneca, Janssen-Cilag, Bayer, Roche Products Ltd, Merck Sharp & Dohm, Accuray Inc. N.v.A. reports research grants from Varian and Accuray, outside of the submitted work. The other authors have nothing to disclose.

Trial Management Group: Isabel Syndikus (chair), Angela Baker, Margaret Bidmead, Ananya Choudhury, David Dearnaley, Vincent Khoo, Helen Mayles, Philip Mayles, Olivia Naismith, Paul Ridley, Christopher Scrase, Chris South, John Staffurth, Alison Tree, Jean Tremlett, Emma Hall, Clare Cruickshank, Clare Griffin, and Julia Pugh.

Data Sharing Statement: Deidentified individual participant data, together with a data dictionary defining each field in the set, will be made available to other researchers on request.

The Institute of Cancer Research (ICR) Clinical Trials and Statistics Unit (CTSU) supports wider dissemination of information from the research it conducts to increase cooperation between investigators. Trial data are obtained, managed, stored, shared, and archived according to ICR-CTSU standard operating procedures to ensure the enduring quality, integrity, and utility of the data. Formal requests for data sharing are considered in line with ICR-CTSU procedures, with due regard given to funder and sponsor guidelines. Requests are via a standard proforma describing the nature of the proposed research and extent of data requirements. Data recipients are required to enter a formal data sharing agreement, which describes the conditions for release and requirements for data transfer, storage, archiving, publication, and Intellectual Property. Requests are reviewed by the Trial Management Group in terms of scientific merit and ethical considerations, including patients' consent. Data sharing is undertaken if proposed projects have a sound scientific or patients' benefit rationale, as agreed by the Trial Management Group and approved by the Independent Data Monitoring and Steering Committee, as required. Restrictions relating to patients' confidentiality and consent will be limited by aggregating and anonymizing identifiable patients' data. Additionally, all indirect identifiers that could lead to deductive disclosures will be removed in line with Cancer Research UK data sharing guidelines.

Acknowledgments—We thank the patients and all investigators and research support staff at the participating centers, Trial Management Group members, the Independent Data Monitoring Committee and Trial Steering Committee. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ijrobp.2022.08.030.

Int J Radiation Oncol Biol Phys, Vol. 000, No. 00, pp. 1–10, 2022

https://doi.org/10.1016/j.ijrobp.2022.08.030

University College London Hospitals NHS Foundation Trust, London, United Kingdom

Received Jun 9, 2022; Accepted for publication Aug 6, 2022

Purpose: Moderately hypofractionated external beam intensity modulated radiation therapy (RT) for prostate cancer is now standard-of-care. Normal tissue toxicity responses to fraction size alteration are nonlinear: the linear-quadratic model is a widely used framework accounting for this, through the α/β ratio. Few α/β ratio estimates exist for human late genitourinary endpoints; here we provide estimates derived from a hypofractionation trial.

Methods and Materials: The CHHiP trial randomized 3216 men with localized prostate cancer 1:1:1 between conventionally fractionated intensity modulated RT (74 Gy/37 fractions (Fr)) and 2 moderately hypofractionated regimens (60 Gy/20 Fr and 57 Gy/19 Fr). RT plan and suitable follow-up assessment was available for 2206 men. Three prospectively assessed clinician-reported toxicity scales were amalgamated for common genitourinary endpoints: dysuria, hematuria, incontinence, reduced flow/stricture, and urine frequency. Per endpoint, only patients with baseline zero toxicity were included. Three models for endpoint grade ≥1 (G1+) and G2+ toxicity were fitted: Lyman Kutcher-Burman (LKB) without equivalent dose in 2 Gy/Fr (EQD2) correction [LKB-NoEQD2]; LKB with EQD2-correction [LKB-EQD2]; LKB-EQD2 with dose-modifying-factor (DMF) inclusion [LKB-EQD2-DMF]. DMFs were age, diabetes, hypertension, pelvic surgery, prior transurethral resection of prostate (TURP), overall treatment time and acute genitourinary toxicity (G2+). Bootstrapping generated 95% confidence intervals and unbiased performance estimates. Models were compared by likelihood ratio test.

Results: The LKB-EQD2 model significantly improved performance over LKB-NoEQD2 for just 3 endpoints: dysuria G1+ (α/β = 2.0 Gy; 95% confidence interval [CI], 1.2-3.2 Gy), hematuria G1+ (α/β = 0.9 Gy; 95% CI, 0.1-2.2 Gy) and hematuria G2+ (α/β = 0.6 Gy; 95% CI, 0.1-1.7 Gy). For these 3 endpoints, further incorporation of 2 DMFs improved on LKB-EQD2: acute genitourinary toxicity and prior TURP (hematuria G1+ only), but α/β ratio estimates remained stable.

Conclusions: Inclusion of EQD2-correction significantly improved model fitting for dysuria and hematuria endpoints, where fitted α/β ratio estimates were low: 0.6 to 2 Gy. This suggests therapeutic gain for clinician-reported GU toxicity, through hypofractionation, might be lower than expected by typical late α/β ratio assumptions of 3 to 5 Gy.

Introduction

Moderately hypofractionated external beam radiation therapy (EBRT) for localized prostate cancer is now standard-of-care.[1-3] This follows the results of 3 major phase 3 studies (CHHiP,[4] PROFIT,[5] and RTOG-0415[6]), each of which confirmed moderate hypofractionation as noninferior for disease control.

Beside the convenience of fewer fractions, hypofractionation for prostate cancer has been of significant interest due to evidence for a low tumor α/β ratio <2 Gy,[7,8] an inverse marker of fraction size sensitivity. There is the potential for therapeutic gain with hypofractionation if the prostate tumor α/β ratio is lower than those of relevant late normal tissue toxicities.[9] It is therefore of interest to have robust estimates of α/β ratios for the common late toxicities after prostate EBRT (eg, gastrointestinal [GI] and genitourinary [GU] toxicities).

Multiple human estimates exist for late GI side effect α/β ratios[10-13]; however, data for GU effects are sparser. Mouse models with an endpoint of reduced bladder capacity have suggested α/β ratios in the range of 3.7 to 5.8 Gy.[14,15] However, limited data from humans have suggested an α/β ratio below 1 Gy provides the best fit to observed clinician[16] and patient reported[17] GU late effects. One of these studies suggested that incorporation of a normal tissue recovery time factor resulted in a stable model with higher α/β ratio fits (3.5-6.5 Gy).[16]

This study reports fitting of α/β ratios for a range of common late GU clinician reported toxicities seen in a large phase 3 hypofractionation trial for localized prostate cancer. Additionally, the effect of including multiple potential dose-modifying factors is examined: age, diabetes, hypertension, pelvic surgery, prior transurethral resection of prostate (TURP), overall treatment time (RT delivery), and the occurrence of acute GU toxicity (up to 18 weeks from RT start date) at RTOG grade 2 or higher (G2+).

Methods and Materials

The CHHiP *trial*

The CHHiP trial (Clinical Trial Number: ISRCTN97182923) has previously been described in detail, including information on written informed consent and ethics approval.[4] In short, it recruited 3216 men with localized prostate adenocarcinoma (T1b-T3a N0 M0, prostate specific antigen ≤40 ng/mL, estimated lymph node risk <30%), randomizing 1:1:1 between 74 Gy in 37 fractions (conventional) over 7.5 weeks, 60 Gy in 20 fractions over 4 weeks or 57 Gy in 19 fractions over 3.8 weeks. Androgen deprivation therapy was given for 3 to 6 months before and during RT (which could be omitted in National Comprehensive Cancer Network low-risk patients). The trial demonstrated noninferiority of the 60 Gy regimen, compared with 74 Gy. Development of bladder normal tissue complication probability (NTCP) models was a protocol-specified objective.

Patient inclusion

All randomized patients were considered for inclusion. Patients were excluded from this NTCP substudy if they received a nonprotocol-defined dose-fractionation regimen, or RT plan information was not available in Digital Imaging and Communications in Medicine (DICOM) format (CT, structures, dose). Efforts were made to convert all non-DICOM treatment plan formats into DICOM. Further patients were excluded on a model-by-model basis, depending on availability of suitable endpoint data.

Bladder dosimetry

Before the planning CT and each delivered fraction, the protocol recommended a "comfortably full" bladder, with around 350 mL fluid ingested, during the hour before scanning. The CHHiP trial protocol defined the bladder as a solid structure, "outlined from base to dome." The bladder DVH was extracted from patients DICOMs, using the bladder structure contoured for treatment planning, as defined by the treating center.

Endpoints

Multiple clinician-reported toxicity scales were assessed during the trial: Radiation Therapy Oncology Group (RTOG) late rectal toxicity,[18] the Royal Marsden Hospital scale,[19] and Late Effects Normal Tissue—Subjective, Objective & Management.[20] The utilized patient reported outcome (PRO) scales varied during the trial, therefore modeling was restricted to clinician-reported outcome (CRO) measures, to maximize patient numbers in any given model. Relevant CRO collection was at baseline and pre-RT (Royal Marsden Hospital and Late Effects Normal Tissue—Subjective, Objective & Management scales only); then 6, 12, 18, 24, 36, 48, and 60 months after RT commencement (all scales). CROs before 6 months were discarded from endpoint generation, given they may represent acute toxicity.

To avoid excessive numbers of models, the CRO scales were merged into unified endpoints representing common GU toxicities: dysuria, hematuria, incontinence, obstructive (reduced flow or stricture), and urine frequency. For this we used a methodology similar to a previously published study,[13] the full process being detailed in Appendix EA. All toxicity was simplified to G0 (no toxicity), G1 (toxicity without intervention), and G2 (toxicity with intervention). Models were then built separately for G1+ and G2+ toxicity. The toxicity score was the worst recorded during the late toxicity follow-up (ie, 6 months post-RT to 5 years). Only patients with documented zero baseline signs or symptoms (G0 toxicity) in an endpoint were included. Patients were only scored as G0 where at least 50% of follow-up assessments were completed (ie, ≥4/7 late toxicity assessments), otherwise they were treated as missing for that endpoint. A summary of patient numbers included in each endpoint along with rates of toxicity is in Table EA. Event rates ranged from 5% (102/2050, hematuria G2+) to 56.7% (371/654, urinary frequency G1+). Patient exclusion rates on a per endpoint basis, ranged from 4.3% (95/2206, dysuria models) to 70.7% (1559/2206 urinary frequency G2+). These were differences largely driven by reported baseline signs or symptoms.

Dose modifying factors

Potential dose-modifying factors (DMFs) were selected from data collected within the trial; age (years), diabetes (yes/no), hypertension (yes/no), pelvic surgery (yes/no), prior TURP (yes/no), treatment days (days from first to last fraction), and acute GU toxicity G2+ (yes/no). Most DMFs had low missing data rates: age (0/2206, 0%), diabetes (18/2206, 0.8%), hypertension (23/2206, 1%), pelvic surgery (17/2206, 0.8%), and prior TURP (43/2206, 2%). For those with missing data, the modal average (DMF absent) was imputed. For one patient where the end date of RT was unobtainable, treatment days was assumed to be 26 days (19 weekday fractions with intervening weekends).

Acute GU toxicity G2+ status was determined from the maximum recorded acute RTOG toxicity, which was collected weekly during weeks 1 to 8, then at 10, 12, and 18 weeks. The DMF acute GU toxicity G2+ was set as missing for patients missing more than half of the assessments; overall missingness was higher than other DMFs (514/2206, 23.3%), so no imputation was used. Instead, separate models were fitted which only included patients who had complete data for this DMF (complete case analysis only).

Modeling

A description of a generalized Lyman-Kutcher Burman (LKB) model has previously been published for an analysis of rectal α/β ratios,[13] derived from prior work by Tucker et al.[21] A full description of all models is described in Appendix EB. In brief, the fitted parameters for the basic model (LKB-NoEQD2 model) were n (measure of organ seriality; nearer zero is more serial); m (related inversely to dose response steepness); and TD50 (the uniform dose required to yield 50% toxicity rate). This was then extended to a model incorporating EQD2 correction for physical dose (LKB-EQD2 model), additionally fitting the α/β ratio (ie, allowing the α/β ratio to vary, to best fit the data). Further models were developed with DMFs incorporated (LKB-EQD2-DMF), separately for each endpoint and DMF.

Model fitting has also previously been described in detail.[13] Briefly, a grid search method was performed for each parameter, with ranking of the best model fits discovered. Assessment of model fit was by sum of the negative log likelihoods. The 10 best grid search positions were then used as the starting point for a constrained Nelder-Mead simplex algorithm[22] search to further optimize the estimates. The best fit after this process was selected as the final model fit.

This process was repeated for all 10 endpoints and 3 models, both for the full population (for that endpoint) and also for 2000 bootstraps, sampled with replacement, for each endpoint modeled. Negative log likelihoods for the best fits for both the full population and bootstrapped fits

were then used to calculate the 632 estimator.[23] This provides a less biased estimate of test performance, where predictions are not close to perfect.

Comparison of nested models could then be made by means of the likelihood ratio test, comparing negative log likelihoods as calculated by 632 estimation. For each independent endpoint comparison of LKB-NoEQD2 to LKB-EQD2 model, a P value of .05 was accepted as significant. Due to multiple testing of DMFs, for the LKB-EQD2 to LKB-EQD2-DMF model comparison, an adjusted P value of .0083 (0.05/6 tests) was deemed significant based on Bonferroni correction.[24] Nonparametric bootstrapped 95% confidence intervals were calculated as the 2.5th to 97.5th percentile bootstrap estimates for all parameters. Calibration plots for each model were calculated, putting patients into 10th percentile bins based on predicted toxicity, then plotting observed vs predicted toxicity for each decile as a point.

Software

All trial data was processed in Stata (version 15/16, Statacorp, TX). VODCA (v5.4.1, Medical Software Solutions GmbH, Switzerland) was used to convert non-DICOM data to DICOM and to check DICOM plan consistency. All models were built in MATLAB (v2018b-v2020a, Mathworks, MA), which was also used for graphical plots. Nelder-Mead simplex algorithm searches were by a modified bounded version of *fminsearch* (fminsearchbnd, v 1.4.0.0).[25]

Results

Patients and endpoints

The CHHiP trial recruited a total of 3216 patients between October 18, 2002, and June 17, 2011. Of these, 2206 patients were able to be included in this substudy, with Figure 1 demonstrating the reasons for noninclusion of all randomized patients. The baseline characteristics of this substudy are shown in Table 1 and are compared with the whole trial population by χ^2 goodness of fit. Only T-stage was significantly different to the whole trial population, although the magnitude of such difference is small.

Non-EQD2 corrected models (LKB-NoEQD2)

The fits of the simplest model, LKB-NoEQD2, are reported in Table EB for 3 groups: 74 Gy in 37 fractions only (2 Gy/fraction), 60 Gy in 20 fractions plus 57 Gy in 19 fractions (3

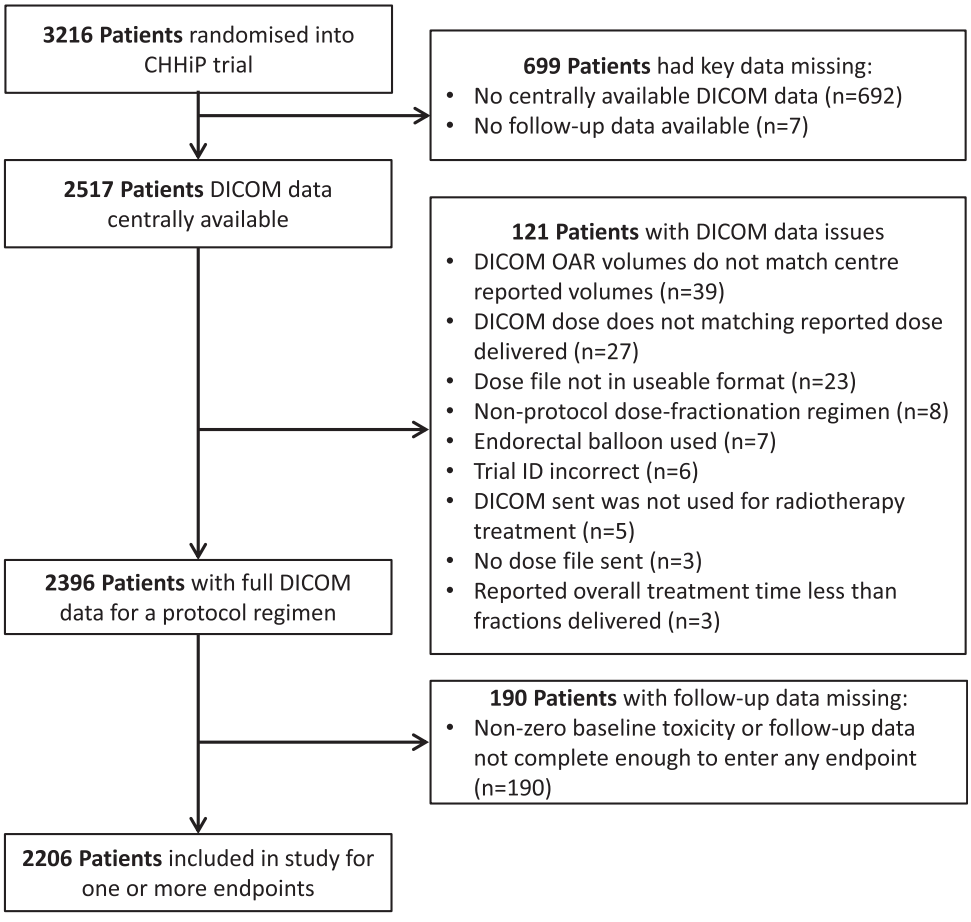

Fig. 1. Patient flow diagram. Showing any reasons for exclusion of all patients originally randomized into the CHHiP trial.

Table 1 Comparison of baseline characteristics between this substudy and whole CHHiP trial population

	Bladder α/β ratio substudy		Whole CHHiP trial		
Characteristic	**Median**	**Range**	**Median**	**Range**	**χ^2 goodness of fit**
Age, y	69	52-80	69	44-85	
Arm	No.	%	No.	%	
57 Gy/19f	751	34	1077	33	
60 Gy/20f	749	34	1074	33	.54
74 Gy/37f	706	32	1065	33	
NCCN risk group					
Low risk	306	14	484	15	
Intermediate risk	1652	75	2347	73	.12
High-risk	248	11	385	12	
Gleason score					
≤6	747	34	1122	35	
7	1394	63	1995	62	.53
8	65	3	99	3	
Clinical T stage					
T1	849	39	1170	36	
T2	1191	54	1766	55	.04
T3	165	8	277	9	
Missing	1	0	3	<1	
Pre-ADT PSA					
<10 ng/mL	1075	49	1567	49	
10-20 ng/mL	1005	46	1415	44	.22
≥20 ng/mL	126	6	208	6	
Missing	0	0	26	<1	
Comorbidities					
Diabetes (yes)	225	10	342	11	.51
Hypertension (yes)	866	39	1276	40	.69
Pelvic Surgery (yes)	165	8	252	8	.53
Prior TURP (yes)	174	8	259	8	.77
Total	2206	100	3216	100	

Note that for comorbidities, only the numbers with that comorbidity present are shown. The other patients are either no or missing, per missing rates reported in the manuscript. χ^2 goodness of fit tests show that only T-stage differs slightly for this sample versus the overall trial population.

Abbreviations: ADT = androgen deprivation therapy; NCCN = National Comprehensive Cancer Network; PSA = prostate specific antigen; TURP = transurethral resection of prostate.

Gy/fraction), and the whole substudy population. The fits are expectedly difficult for the whole study population (because no EQD2 correction is applied to physical dose), with 95% CI for many model parameters restrained at the limits of the searched space.

EQD2 corrected models (LKB-EQD2)

The fits of the next model, LKB-EQD2, incorporating α/β ratio correction, are presented in Table 2. The LKB-EQD2 model fit was significantly better than the LKB-NoEQD2 model for 3 of the endpoints examined: dysuria G1+ (P = .0046), hematuria G1+ (P = .034), and hematuria G2+ (P = .015). Further fitting of the other 7 endpoints was stopped at this stage. For the 3 endpoints with significant improvement, we see that the fitted α/β ratio was low and ranged from 0.6 to 2.0 Gy, with the 95% confidence intervals for these 3 estimates encompassed by 0.1 to 3.2 Gy. Bootstrap distributions of the fitted α/β ratios in these 3 models are shown in Figure E A-C. An example decile bin calibration plot for hematuria G2+ is shown in Figure 2. Similar calibration plots for dysuria G1+ and hematuria G1+ are shown in respectively Figure E D and E.

Table 2 LKB-EQD2 model fits for all endpoints

LKB-EQD2 (all pts)	Patients	n	m	TD50	α/β ratio	632 likelihood	Likelihood ratio test *P* value
Dysuria G1+	2111	0.02 (0.01-0.07)	0.19 (0.13-0.37)	89.9 (82.5-118.4)	2.0 (1.2-3.2)	−793.7	.0046
Dysuria G2+	2111	0.02 (0.01-1.00)	0.26 (0.15-0.59)	119.7 (93.3-594.3)	1.6 (0.1-36.0)	−75.5	Worse fit
Haematuria G1+	2053	0.07 (0.02-0.40)	0.32 (0.19-0.58)	110.3 (89.4-184.1)	0.9 (0.1-2.2)	−673.2	.034
Hematuria G2+	2050	0.04 (0.01-0.12)	0.24 (0.14-0.40)	120.1 (93.6-204.5)	0.6 (0.1-1.7)	−403.1	.015
Incontinence G1+	1927	0.01 (0.01-1.26)	0.55 (0.26-1.27)	125.6 (89.5-617.0)	1.0 (0.1-17.6)	−1051.7	Worse fit
Incontinence G2+	1923	0.02 (0.01-0.23)	0.24 (0.14-0.59)	108.9 (88.8-270.6)	1.5 (0.1-6.2)	−494	.24
Reduced flow/stricture G1+	1743	0.17 (0.06-0.45)	0.72 (0.38-1.19)	95.0 (76.0-182.2)	1.9 (0.1-424.6)	−995.4	.35
Reduced flow/stricture G2+	1743	0.17 (0.01-10.00)	0.58 (0.27-0.82)	160.0 (98.1-686.0)	0.7 (0.1-991.9)	−656.1	Worse fit
Urine frequency G1+	654	0.60 (0.01-4.21)	5.96 (0.27-10.00)	15.3 (6.6-67.3)	1.9 (0.1-997.8)	−447.3	Worse fit
Urine frequency G2+	647	0.02 (0.01-0.45)	0.32 (0.15-1.03)	90.4 (76.1-263.7)	3.3 (0.1-996.0)	−334.9	Worse fit

Fits of the Lyman-Kutcher-Burman model with equivalent dose in 2 Gy/fraction correction (LKB-EQD2) model for all endpoints. The parameter estimates are those derived from the naïve model. Bootstrapped 95% percentile confidence intervals are presented for the model parameter estimates. The likelihood ratio *P* value test highlights those models where the LKB-EQD2 model significantly outperformed the non-EQD2 corrected LKB-NoEQD2 model (fitted to whole substudy population) of the same endpoint. Worse fits are possible with this more complex model due to the 632-likelihood penalizing overfitting.

Dose modifying factor fits (LKB-EQD2-DMF)

More complex models incorporating DMFs were then fitted for the 3 endpoints where EQD2 correction significantly improved the model: dysuria G1+, hematuria G1+, and hematuria G2+. These fits are presented in Table E C, with the LKB-EQD2 model fit again shown at the top of each endpoint section to aid comparison. Only one LKB-EQD2-DMF model met *P* value threshold for a significant improvement over LKB-EQD2 alone, hematuria G1+ with prior *TURP* (yes/no) as a DMF ($P = .0015$). The α/β ratio estimate was similar to the hematuria G1+ LKB-EQD2 model without DMF correction (both 0.9 Gy). The calibration plot for this model is shown in Figure E F. A signal for prior TURP also being beneficial to the hematuria G2+ model is also seen ($P = .02$) but did not meet significance after adjusting for multiple comparisons.

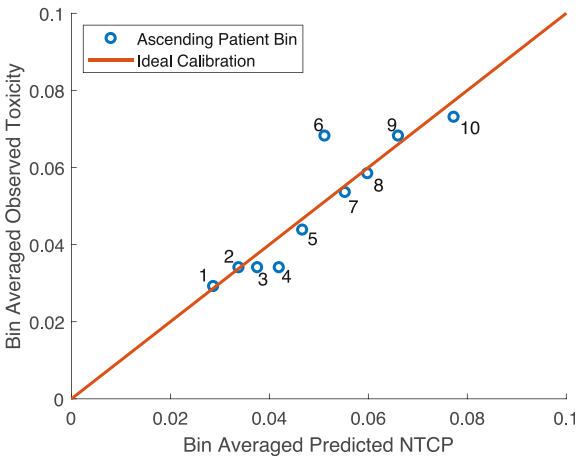

Fig. 2. Calibration plot for LKB-EQD2 hematuria G2+ model. Binned decile calibration plot for hematuria G2+, one endpoint where the LKB-NoEQD2 model were significantly improved by EQD2 correction (LKB-EQD2). Bins are by decile of predicted NTCP, plotted on the x-axis versus observed toxicity on the y-axis. Perfect prediction is shown as the orange identity line. *Abbreviations*: LKB-EQD2 = Lyman-Kutcher-Burman model with equivalent dose in 2 Gy/fraction correction.

The fitted α/β ratios in these LKB-EQD2-DMF models are generally similar. A slight exception is the treatment days DMF, where although parameter estimates of the α/β ratios remained low across the 3 endpoints (0.1-1.2 Gy), the 95% confidence intervals were greatly increased for dysuria G1+ and hematuria G1+. All of the LKB-EQD2-DMF models incorporating treatment days as a DMF performed worse by 632 likelihood than the simpler LKB-EQD2 model, suggesting overfitting.

Acute toxicity as a DMF

As outlined previously, in examining acute toxicity as a DMF (of particular interest due to possibility of consequential late effects), a complete-case approach was adopted due to significant missing data on acute toxicity. Therefore, for acute GU toxicity G2+, separate LKB-NoEQD2, LKB-EQD2, and LKB-EQD2-DMF models were fitted for dysuria G1+, hematuria G1+ and hematuria G2+, including only patients with recorded data (yes/no) for acute GU toxicity G2+. The fits for these models are presented in Table 3. For hematuria G1+, it can be seen that in this subset, LKB-EQD2 model did not significantly outperform LKB-NoEQD2. However, we proceeded

Table 3 Acute toxicity GU G2+ as dose modifying factor

Model	n	m	TD50	α/β ratio	Dose-modifying factor	632 likelihood	Likelihood ratio test *P* value
Dysuria G1+ (n = 1611)							
LKB-NoEQD2	.48 (0.01-1.86)	.76 (0.56-0.85)	253.6 (119.3-1000.0)	0 (0.0-0.0)	N/A	−614.5	
LKB-EQD2	.02 (0.01-0.11)	.19 (0.12-0.46)	89.9 (81.8-132.6)	1.6 (0.5-2.5)	N/A	−610.4	.0044
LKB-EQD2-DMF	.02 (0.01-0.78)	.19 (0.12-0.71)	97.6 (85.5-999.9)	1.5 (0.1-2.4)	1.11 (1.05-10.92)	−599.1	2.00E-06
Hematuria G1+ (n = 1576)							
LKB-NoEQD2	1.93 (0.04-10.00)	.75 (0.67-0.78)	550.4 (166.8-1000)	0 (0.0-0.0)	N/A	−524.5	
LKB-EQD2	.06 (0.01-4.76)	.41 (0.23-0.73)	141.7 (98.6-346.3)	.1 (0.1-2.1)	N/A	−523.6	.16
LKB-EQD2-DMF	.05 (0.01-0.37)	.38 (0.23-0.66)	158.4 (106.1-999.8)	.4 (0.1-262.0)	1.27 (1.09-5.63)	−517.6	.00054
Hematuria G2+ (n = 1574)							
LKB-NoEQD2	8.64 (0.01-10.00)	.59 (0.56-0.59)	1000 (274.3-1000)	0 (0.0-0.0)	N/A	−311	
LKB-EQD2	.02 (0.01-0.11)	.29 (0.18-0.50)	148.9 (106.8-358.3)	.1 (0.1-1.4)	N/A	−309	.048
LKB-EQD2-DMF	.01 (0.01-0.10)	.29 (0.18-0.48)	170.7 (115.6-799.0)	.1 (0.1-1.3)	1.22 (1.06-2.77)	−305	.0046

Fits of the LKB-NoEQD2 model, LKB-EQD2 model, and LKB-EQD2-DMF model (acute toxicity GU G2+) for patients without missing data for acute toxicity GU G2+. Shown for those endpoints where in main analysis EQD2 correction significantly improved the LKB-LKB-NoEQD2 model: dysuria G1+, hematuria G1+, hematuria G2+. The parameter estimates are those derived from the naïve model, with 95% confidence intervals from bootstrapping percentiles. Inclusion of acute toxicity always significantly improves the model but does not appear to strongly increase the fitted α/β ratios. This suggests a limited detectable effect from consequential late effects arising from high α/β ratio acute reactions.

to fit the LKB-EQD2-DMF model based on the significant improvement seen earlier at the whole trial level (Table 2). The LKB-EQD2-DMF (acute GU toxicity G2+) was significantly better than LKB-EQD2 model for all endpoints. For dysuria G1+ and hematuria G2+, the model parameters, including α/β ratios were highly stable. For hematuria G1+, the inclusion of DMF = acute GU toxicity G2+ resulted in a very slightly higher α/β ratio estimate, with much wider bootstrapped 95% confidence interval: LKB-EQD2 α/β 0.1 Gy (95% CI, 0.1-2.1 Gy) versus LKB-EQD2-DMF α/β 0.4 Gy (95% CI, 0.1-262 Gy).

Discussion

In this study, we aimed to provide human estimates of GU fraction size sensitivity across several common toxicity endpoints after prostate EBRT. We found only 3 endpoints had improved models when EQD2 correction was incorporated: dysuria G1+, hematuria G1+, and hematuria G2+. The α/β ratio estimates in these models were low, ranging from 0.6 to 2.0 Gy. The failure of most models to benefit from EQD2-adjustment suggests dose-toxicity relationships for those endpoints are weak. When incorporating DMFs, other than for acute toxicity GU G2+, only one model met adjusted P value significance for improvement over an LKB-EQD2 model without DMF correction. This was hematuria G1+, with TURP as DMF; however, the α/β ratio estimate remained low (0.9 Gy; 95% CI, 0.1-2.6).

Inclusion of treatment days as a DMF did not improve model fits, though nor did it greatly alter the α/β ratio estimates. This is important given that the CHHiP trial delivered hypofractionated treatment over a reduced overall treatment time in the investigational arms. Significant normal tissue repopulation effects (protecting from toxicity) might alter the apparent α/β ratio for late effects.

Including acute GU toxicity G2+ as a DMF improved model fits, with very limited effect on α/β ratio estimates, indicating minimal contribution from consequential late effects. This is important because if substantial consequential late effects were present (stemming from high α/β ratio acute effects) then this may raise the apparent α/β ratio of late effects as a whole.

Prior estimates of α/β ratios for bladder endpoints

Fiorino retrospectively examined a single center cohort of men receiving postprostatectomy RT: either conventionally fractionated (n = 929, 1.8 Gy/fraction, dose 60-77.4 Gy) or hypofractionated (n = 247, 2.35-2.90 Gy/fraction, dose 58-71.4 Gy).[16] Modeling of Common Terminology Criteria for Adverse Events endpoints was by logit function, fitting a TD50 and slope: seriality was not fitted. Fitting without a time factor, they estimated the α/β ratio as 0.81 Gy (95% CI, 0.1-4.8) for G3+ urinary incontinence and 0.74 Gy (95% CI, 0.0-4.8) for G3+ hematuria. They suggested that inclusion of a time factor with a fixed α/β ratio of 5 Gy resulted in reasonable models, although direct comparison of fits was not reported. Simultaneous fitting of α/β ratio with a time factor had wide confidence intervals for α/β ratios (95% CI, ≈0.4-10 Gy, almost the entire searched space). This study has limitations from a population perspective (retrospectively graded endpoints) and a modeling perspective (incomplete dosimetric information, no seriality in model).

The DUE01 study prospectively collected toxicity data on patients treated with primary EBRT for PCa, either conventionally fractionated or moderately hypofractionated. Cozzarini et al reported fitting of multivariate logistic models to the International Consultation on Incontinence Questionnaire Short Form (ICIQ-SF) urinary incontinence questionnaire.[17] They adjusted the prescription doses entered into the model by EQD2, testing α/β ratio = 0.8 Gy, α/β ratio = 3 Gy, and α/β ratio = 5 Gy. They found that the α/β ratio of 0.8 Gy provided the best fit for the data.

Overall, if we accept that inclusion of a time factor does not improve model fitting, then our results are largely similar to these 2 prior studies, with all reported α/β parameter estimates being <3 Gy. We did not find that EQD2 adjusted urinary incontinence models statistically improved from non-EQD2 adjusted models.

EQD2 correction failing to improve LKB model for most endpoints

Adding EQD correction (LKB-EQD2) to the LKB model (LKB-NoEQD2) failed to improve model fits for dysuria G1+, incontinence (G1+/G2+), reduced flow/stricture (G1+/G2+), urine frequency (G1+/G2+). We would expect that this is due to the relatively weak relationship between dose to whole bladder and subsequent toxicity in a genitourinary setting. Many patients with prostate cancer are elderly, with a substantial chance of developing lower urinary tract symptoms over 5 years in the absence of prostate cancer treatment. This may add substantial nondose-related toxicity noise to the model, even with relatively stringent selection criteria, such as requiring baseline zero GU toxicity.

Late genitourinary toxicity in phase 3 hypofractionation trials

Five major phase 3 hypofractionation trials have reported cumulative 5-year genitourinary toxicity outcomes, with one announcing 2-year data.[4-6,26-28] These are summarized in Table 4. Of most interest to this analysis are trials that are close to isotoxic for GU late effects: PROFIT[5] and HYPO-RT-PC.[27] Both had identical RT planning and delivery methods for the conventional and hypofractionated arms. The α/β ratio for isotoxicity in PROFIT (α/β = 1.3 Gy) and HYPO-RT-PC (α/β = 2.9 Gy) both lie within the 95% confidence interval for the GU toxicities we have fitted in this article. It is worth noting the range of α/β ratio assumptions made across the trials, with all assumptions being higher than the estimated α/β ratios we have found in this study. Our results imply that any potential GU toxicity therapeutic gain from isoeffective hypofractionation might be lower than predicted at the time of study design.

Comparison with gastrointestinal late α/β ratio estimates

The estimates for late GU α/β ratios in this study, along with those seen in prior human studies,[16,17] are all ≤2 Gy. This contrasts with higher estimates generally seen for late GI toxicity, with published individual patient level estimates

Table 4 Summary of hypofractionation trials and genitourinary toxicity

Trial [reference]	Dose- fractionation schedule (PTV doses)		Assumed late effects α/β ratio	Toxicity scale	Cumulative late genitourinary toxicity			
	Control	HypoFr			G2+		G3+	
					Control	HypoFr	Control	HypoFr
CHHiP [4]	74 Gy in 37 Fr	60 Gy in 20 Fr 57 Gy in 19 Fr	3 Gy	RTOG	9%	12% 7%	3%	6% 3%
HYPRO [26]	78 Gy in 39 Fr	64.6 Gy in 19 Fr	4-6 Gy	RTOG	39%	41%	13%	19%
PROFIT [5]	78 Gy in 39 Fr	60 Gy in 20 Fr	3-5 Gy	RTOG	22%	22%	3%	2%
RTOG 0415 [6]	73.8 Gy in 41Fr	70 Gy in 28 Fr	3 Gy	CTCAE	23%	30%	2%	3%
HYPO-RT-PC [27]	78 Gy in 39 Fr	42.7 Gy in 7 Fr	3 Gy	RTOG	17%	18%	5%	4%
PACE-B [28]	78 Gy in 39 Fr 62 Gr in 20 Fr	36.25 Gy in 5 Fr*	3 Gy	RTOG	13%	21%	N/A	N/A
				CTCAE	19%	29%	N/A	N/A

Table showing the 6 major phase 3 trials of hypofractionated external beam radiation therapy schedules for prostate cancer. Reported cumulative rates of GU toxicity are shown. Cumulative late toxicity shown at 5 years, except PACE-B where 2-year cumulative toxicity is shown (G3 data yet to be published).

* Note that PACE-B had a secondary dose level of 40 Gy to the CTV.

Abbreviations: CTCAE = common terminology criteria for adverse events; Fr = Fractions; GX+ = grade X or more; GU = genitourinary; HypoFr = hypofractionated; N/A = not available; PTV = planning target volume; RTOG = Radiation Therapy Oncology Group.

ranging 2.3 to 4.8 Gy.[11-13] This suggests that late GU side effects might have a greater hypofractionation sensitivity than late GI side effects. This would be supported by meta-analysis of moderately hypofractionated versus conventional RT noninferiority trials, which noted an increase with hypofractionation for late GU G2+ toxicity (relative risk 1.18; 95% CI, 0.98-1.43; $P = .08$), but no significant increase for late GI G2+ toxicity.[29] It would also be supported by the higher late toxicity data in PACE-B, where 2-year incidence of Common Terminology Criteria for Adverse Events late toxicity was worse with ultrahypofractionation (36.25 Gy in 5 fractions) versus conventional or moderately hypofractionated RT (78 Gy in 39 fractions or 62 Gy in 20 fractions) for GU (29% vs 19%), and GI was similar (12% vs 10%).[28] Clinician reported data from HYPO-RT-PC do not appear to support this, with RTOG 5-year cumulative late toxicity similar for ultrahypofractionation (42.7 Gy in 7 fractions) versus conventional (78 Gy in 39 fractions), both for GU (18% vs 17%) and GI (6% vs 5%).[27] There was also no long-term differences in urinary quality-of-life outcomes in HYPO-RT-PC,[30] including pain on urinating (approximately to the dysuria modelled in our study), although we note that hematuria is not included on the PCSS (Prostate Cancer Symptom Scale) questionnaire. As an aside, it is interesting that we see similar α/β ratio estimates for late hematuria G1+ here (α/β 0.9 Gy; 95% CI, 0.1-2.2 Gy) and late rectal bleeding G1+ in our prior work (α/β 1.6 Gy; 95% CI, 0.9-2.5 Gy),[13] potentially due to similar biological pathways.

Strengths

There are several strengths of this study related to its data source. This is, to date, the largest reported cohort where α/β ratio estimates have been made for late GU toxicity. Data were prospectively collected and have generally low levels of missingness for model covariates. The modeling methodology has used bootstrapping to avoid overfitting. We have observed stability of reported α/β ratio estimates with inclusion of multiple DMFs, including overall treatment time and acute toxicity, both of which have a strong rationale to potentially alter these estimates.

Limitations

There is a strong overlap between genitourinary toxicity and lower urinary tract symptoms, which might also occur in men typically treated for prostate cancer. The endpoints are clinician reported, which may differ to patient reported metrics.[31] Additionally, we have examined cumulative toxicity, which does not account for any recovery seen after RT[4]; this was chosen to avoid missing patients whose score recovered due to management of the toxicity. Although the LKB model is a well-studied method of examining dose-response relationships, it is limited in terms of incorporating the multitude of potentially causative factors for lower urinary tract symptoms. Another limitation is the relatively narrow range of doses included in the study (2-3 Gy/fraction), meaning that caution would be needed in extrapolation to ultrahypofractionation studies. The whole bladder was utilized for dose extraction, which could be a limitation if dose to other putative structures, such as bladder trigone or urethra,[32] are more relevant. The doses examined are planned doses, rather than delivered doses, which have been shown to be better for rectal toxicity prediction,[33] although the role for bladder toxicity is less clear. We also note that population α/β ratio estimates do not account for interpatient heterogeneity that likely exists for normal tissue fraction size sensitivity.

Future work

Given the limited published evidence, it would be desirable for other groups to consider analyzing apparent GU α/β ratios. This could be performed on data arising from other large prostate hypofractionation trials. In particular, individual patient level analysis of data from ultrahypofractionated studies (eg, HYPO-RT-PC and PACE) would be helpful, to confirm applicability at higher doses per fraction.

Conclusion

We have fitted modified Lyman Kutcher-Burman models for individual genitourinary toxicity endpoints after conventional and moderately hypofractionated external beam RT. Three models were statistically improved by use of an EQD2 correction: dysuria G1+, hematuria G1+, and hematuria G2+. For these, the fitted α/β ratio was low, ranging from 0.6 to 2.0 Gy, with 95% confidence intervals for these estimates contained within 0.1 to 3.2 Gy. These low estimates suggest that the therapeutic gain with hypofractionation may be less than expected from modeling using the usual assumption that the late α/β ratio is 3 to 5 Gy. This is important given current trends toward more profoundly hypofractionated RT. For ultrahypofractionated RT in 5 fractions or fewer, technical approaches to reduce genitourinary tract dose may be needed to avoid increasing GU side effects.

References

1. Mottet N, van den Bergh RCN, Briers E, et al. EAU-EANM-ESTRO-ESUR-SIOG guidelines on prostate cancer—2020 update. Part 1: Screening, diagnosis, and local treatment with curative intent. *Eur Urol* 2021;79:243–262.
2. Morgan SC, Hoffman K, Loblaw DA, et al. Hypofractionated radiation therapy for localized prostate cancer: Executive summary of an ASTRO, ASCO, and AUA evidence-based guideline. *Pract Radiat Oncol* 2018;8:354–360.
3. NICE. Prostate cancer: Diagnosis and management. NICE guideline [NG131]. 2019. Available at: https://www.nice.org.uk/guidance/NG131. Accessed June 1, 2020.
4. Dearnaley D, Syndikus I, Mossop H, et al. Conventional versus hypofractionated high-dose intensity-modulated radiotherapy for prostate cancer: 5-year outcomes of the randomised, non-inferiority, phase 3 CHHiP trial. *Lancet Oncol* 2016;17:1047–1060.

5. Catton C, Catton CN, Lukka H, Gu CS, et al. Randomized trial of a hypofractionated radiation regimen for the treatment of localized prostate cancer. *J Clin Oncol* 2017;35:1884–1890.
6. Lee WR, Dignam JJ, Amin MB, et al. Randomized phase III noninferiority study comparing two radiotherapy fractionation schedules in patients with low-risk prostate cancer. *J Clin Oncol* 2016;34:2325–2332.
7. Fowler J, Chappell R, Ritter M. Is α/β for prostate tumors really low? *Int J Radiat Oncol Biol Phys* 2001;50:1021–1031.
8. Vogelius IR, Bentzen SM. Dose response and fractionation sensitivity of prostate cancer after external beam radiation therapy: A meta-analysis of randomized trials. *Int J Radiat Oncol Biol Phys* 2018;100:858–865.
9. Dearnaley D, Hall E. How will the CHHiP trial affect the future of prostate radiotherapy? *Expert Rev Anticancer Ther* 2018;18:607–609.
10. Brenner DJ. Fractionation and late rectal toxicity. *Int J Radiat Oncol* 2004;60:1013–1015.
11. Marzi S, Saracino B, Petrogari MG, et al. Modeling of alpha/beta for late rectal toxicity from a randomized phase II study: Conventional versus hypofractionated scheme for localized prostate cancer. *J Exp Clin Cancer Res* 2009;28:117.
12. Tucker SL, Thames HD, Michalski JM, et al. Estimation of α/β for late rectal toxicity based on RTOG 94-06. *Int J Radiat Oncol Biol Phys* 2011;81:600–605.
13. Brand DH, Brüningk SC, Wilkins A, et al. Estimates of alpha/beta (α/β) ratios for individual late rectal toxicity endpoints: An analysis of the CHHiP trial. *Int J Radiat Oncol Biol Phys* 2021;110:596–608.
14. Dörr W, Bentzen SM. Late functional response of mouse urinary bladder to fractionated X-irradiation. *Int J Radiat Biol* 1999;75:307–315.
15. Bentzen SM, Lundbeck F, Christensen LL, Overgaard J. Fractionation sensitivity and latency of late radiation injury to the mouse urinary bladder. *Radiother Oncol* 1992;25:301–307.
16. Fiorino C, Cozzarini C, Rancati T, et al. Modelling the impact of fractionation on late urinary toxicity after postprostatectomy radiation therapy. *Int J Radiat Oncol Biol Phys* 2014;90:1250–1257.
17. Cozzarini C, Rancati T, Palorini F, et al. Patient-reported urinary incontinence after radiotherapy for prostate cancer: Quantifying the dose-effect. *Radiother Oncol* 2017;125:101–106.
18. Cox JD, Stetz J, Pajak TF. Toxicity criteria of the radiation therapy oncology group (RTOG) and the European Organization for Research and Treatment of Cancer (EORTC). *Int J Radiat Oncol Biol Phys* 1995;3:1334–1346.
19. Dearnaley DP, Sydes MR, Langley RE, et al. The early toxicity of escalated versus standard dose conformal radiotherapy with neo-adjuvant androgen suppression for patients with localised prostate cancer: Results from the MRC RT01 trial (ISRCTN47772397). *Radiother Oncol* 2007;83:31–41.
20. LENT SOMA tables. *Radiother Oncol* 1995;35:1–81.
21. Tucker SL, Liu HH, Liao Z, et al. Analysis of radiation pneumonitis risk using a generalized Lyman Model. *Int J Radiat Oncol Biol Phys* 2008;72:568–574.
22. Nelder JA, Mead R. A simplex method for function minimization. *Comput J* 1965;7:308–313.
23. Efron B, Tibshirani R. Improvements on cross-validation: The .632+ bootstrap method. *J Am Stat Assoc* 1997;92:548–560.
24. Bonferroni CE. Teoria statistica delle classi e calcolo delle probabilità. (1936).
25. D'Errico J. fminsearchbnd, fminsearchcon. (2012). Available at: https://uk.mathworks.com/matlabcentral/fileexchange/8277-fminsearchbnd-fminsearchcon. Accessed January 8, 2020.
26. Incrocci L, Wortel RC, Alemayehu WG, et al. Hypofractionated versus conventionally fractionated radiotherapy for patients with localised prostate cancer (HYPRO): Final efficacy results from a randomised, multicentre, open-label, phase 3 trial. *Lancet Oncol* 2016;17:1061–1069.
27. Widmark A, Gunnlaugsson A, Beckman L, et al. Ultra-hypofractionated versus conventionally fractionated radiotherapy for prostate cancer: 5-year outcomes of the HYPO-RT-PC randomised, non-inferiority, phase 3 trial. *Lancet* 2019;394:385–395.
28. Tree AC, Hall E, Ostler E, et al. Comparison of side effects at 2 years in the randomised PACE-B trial (SBRT vs standard radiotherapy); *ESTRO 2021: Presentation Number: OC-0289*. https://www.estro.org/Congresses/ESTRO-2021/499/profferedpapers16-late-breakingabstracts/3416/comparisonofsideeffectsat2yearsintherandomisedpace. Accessed September 14, 2022.
29. Royce TJ, Lee DH, Keum N, et al. Conventional versus hypofractionated radiation therapy for localized prostate cancer: A meta-analysis of randomized noninferiority trials. *Eur Urol Focus* 2019;5:577–584.
30. Fransson P, Nilsson P, Gunnlaugsson A, et al. Ultra-hypofractionated versus conventionally fractionated radiotherapy for prostate cancer (HYPO-RT-PC): Patient-reported quality-of-life outcomes of a randomised, controlled, non-inferiority, phase 3 trial. *Lancet Oncol* 2021;22:235–245.
31. Rammant E, Ost P, Swimberghe M, et al. Patient- versus physician-reported outcomes in prostate cancer patients receiving hypofractionated radiotherapy within a randomized controlled trial. *Strahlentherapie Onkol* 2019;195:393–401.
32. Marcello M, Denham JW, Kennedy A, et al. Increased dose to organs in urinary tract associates with measures of genitourinary toxicity in pooled voxel-based analysis of 3 randomized phase III trials. *Front Oncol* 2020;10.
33. Shelley LEA, Scaofe JE, Romanchikova M, et al. Delivered dose can be a better predictor of rectal toxicity than planned dose in prostate radiotherapy. *Radiother Oncol* 2017;123:466–471.