Title

The Fraction Size Sensitivity of Late Genitourinary Toxicity: Analysis of alpha/beta (α/β) ratios in the XXXXXX Trial

Abstract

Purpose

Moderately hypofractionated external beam intensity-modulated radiotherapy (IMRT) for prostate cancer is now standard-of-care. Normal tissue toxicity responses to fraction size alteration are non-linear: 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 XXXXXX trial randomised 3216 men with localised prostate cancer 1:1:1 between conventionally fractionated IMRT (74Gy/37 fractions (Fr)) and two moderately hypofractionated regimens (60Gy/20Fr & 57Gy/19Fr). Radiotherapy 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, Haematuria, Incontinence, Reduced flow/Stricture, 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 2Gy/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 three endpoints: Dysuria G1+ (α/β=2.0 Gy, 95%CI 1.2–3.2Gy), Haematuria G1+ (α/β=0.9 Gy, 95%CI 0.1–2.2Gy) and Haematuria G2+ (α/β=0.6Gy, 95%CI 0.1–1.7Gy). For these three endpoints, further incorporation of two DMFs improved on LKB-EQD2: *acute genitourinary toxicity* and *Prior TURP* (Haematuria G1+ only), but α/β ratio estimates remained stable.

Conclusions

Inclusion of EQD2-correction significantly improved model fitting for Dysuria and Haematuria endpoints, where fitted α/β ratio estimates were low: 0.6–2 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 3–5 Gy.

Introduction

Moderately hypofractionated external beam radiotherapy (EBRT) for localised prostate cancer is now standard-of-care.[1–3] This follows the results of three major phase III studies (XXXXXX [4], PROFIT [5] and RTOG-0415 [6]), each of which confirmed moderate hypofractionation as non-inferior for disease control.

Besides the convenience of fewer fractions, hypofractionation for prostate cancer has been of significant interest due to evidence for a low tumour 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 tumour 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 following prostate EBRT; e.g. 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 is more sparse. 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 the limited data from humans has 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 III hypofractionation trial for localised 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 (radiotherapy delivery) and the occurrence of acute GU toxicity (Up to 18 weeks from radiotherapy start date) at RTOG grade 2 or higher (G2+).

Methods and Materials

The XXXXXX *Trial*

The XXXXXX trial (XXXTRIAL_NUMBERXXX) has previously been described in detail, including information on written informed consent and ethics approval.[4] In short, it recruited 3216 men with localised prostate adenocarcinoma (T1b –T3a N0 M0, PSA ≤ 40 ng/mL, estimated lymph node risk <30%), randomising 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 (ADT) was given for 3-6 months before and during radiotherapy (which could be omitted in NCCN low-risk patients). The trial demonstrated non-inferiority of the 60 Gy regimen, compared to 74 Gy. Development of bladder normal tissue complication probability (NTCP) models was a protocol-specified objective.

Patient Inclusion

All randomised patients were considered for inclusion. Patients were excluded from this Normal Tissue Complication Probability (NTCP) sub-study if they received a non-protocol-defined dose-fractionation regimen, or radiotherapy plan information was not available in 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, as described below.

Bladder Dosimetry

Prior to the planning CT and each delivered fraction, the protocol recommended a "comfortably full" bladder, with around 350mL fluid ingested, during the hour before scanning. The XXXXXX 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 centre.

Endpoints

Multiple clinician-reported toxicity scales were assessed during the trial: Radiation Therapy Oncology Group (RTOG) late rectal toxicity [18], the Royal Marsden Hospital (RMH) scale [19] and Late Effects Normal Tissue – Subjective, Objective & Management (LENT-SOM) [20]. The utilised patient reported outcome (PRO) scales varied during the trial, therefore modelling was restricted to clinician-reported outcome (CRO) measures, in order to maximise patient numbers in any given model. Relevant CRO collection was at baseline and pre-radiotherapy (RMH and LENT-SOM scales only); then 6, 12, 18, 24, 36, 48 & 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, haematuria, incontinence,

obstructive (reduced flow or stricture), urine frequency. For this we used a methodology similar to a previously published study [13], the full process being detailed in **Supplementary Appendix A**. All toxicity was simplified to G0 (no toxicity), G1 (toxicity without intervention), 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 (i.e. 6 months post-radiotherapy 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 (i.e. ≥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 **Supplementary Table A**. Event rates ranged from 5% (102/2050, Haematuria 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), *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%), *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 radiotherapy 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 1-8, then at 10, 12, 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 that had complete data for this DMF (complete case analysis only).

Modelling

A description of a generalised 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 **Supplementary Appendix B**. 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); *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 (i.e. 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 ten best grid search positions were then used as the starting point for a constrained Nelder-Mead simplex algorithm [22] search to further optimise 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 modelled. 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 0.05 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 0.0083 (0.05 / 6 tests) was deemed significant based on Bonferroni correction.[24] Non-parametric 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, USA). 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, USA), 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 XXXXXX trial recruited a total of 3216 patients between 18th October 2002, and 17th June 2011. Of these, 2206 patients were able to be included in this sub-study, with **Figure 1** demonstrating the reasons for non-inclusion of all randomised patients. The baseline characteristics of this sub-study are shown in **Table 1** and are compared to the whole trial population by chi-square 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 **Supplementary Table B** for three groups: 74 Gy in 37 fractions only (2 Gy / fraction), 60 Gy in 20 fractions plus 57 Gy in 19 fractions (3 Gy / fraction) and the whole sub-study population. The fits are expectedly difficult for the whole study population (since 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 three of the endpoints examined: Dysuria G1+ (**p=0.0046**), Haematuria G1+ (**p=0.034**) and Haematuria G2+ (**p=0.015**). Further fitting of the other seven endpoints was stopped at this stage. For the three endpoints with significant improvement, we see that the fitted α/β ratio was low and ranged from 0.6 – 2.0 Gy, with the 95% confidence intervals for these three estimates encompassed by 0.1 – 3.2 Gy. Bootstrap distributions of the fitted α/β ratios in these three models are shown in **Supplementary Figures A-C**. An example decile bin calibration plot for Haematuria G2+ is shown in **Figure 2**. Similar calibration plots for Dysuria G1+ & Haematuria G1+ are shown in respectively **Supplementary Figure D and E**.

Dose Modifying Factor Fits (LKB-EQD2-DMF)

More complex models incorporating DMFs were then fitted for the three endpoints where EQD2 correction significantly improved the model: Dysuria G1+, Haematuria G1+ and Haematuria G2+. These fits are presented in **Supplementary Table 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, Haematuria G1+ with *Prior TURP* (yes/no) as a DMF (**p=0.0015**). The α/β ratio estimate was similar to the Haematuria G1+ LKB-EQD2 model without DMF correction (both 0.9 Gy). The calibration plot for this model is shown in **Supplementary Figure F**. A signal for *Prior TURP* also being

beneficial to the Haematuria G2+ model is also seen (p=0.02), but did not meet significance after adjusting for multiple comparisons.

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 three endpoints (0.1 – 1.2 Gy), the 95% confidence intervals were greatly increased for Dysuria G1+ and Haematuria 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 above, 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+, Haematuria G1+ and Haematuria 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 Haematuria G1+, it can be seen that in this subset, LKB-EQD2 model did not significantly outperform LKB-NoEQD2. However, we proceeded to fit the LKB-EQD2-DMF model based on the significant improvement seen earlier at the whole trial level (per **Table 2**). The LKB-EQD2-DMF (*acute GU toxicity G2+*) was significantly better than LKB-EQD2 model for all endpoints. For Dysuria G1+ and Haematuria G2+, the model parameters, including α/β ratios were highly stable. For Haematuria 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) vs 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 a number of common toxicity endpoints following prostate EBRT. We found only three endpoints had improved models when EQD2 correction was incorporated: Dysuria G1+, Haematuria G1+ and Haematuria G2+. The α/β ratio estimates in these models were low, ranging from 0.6 – 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 Haematuria 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 XXXXXX 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, since 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 centre cohort of men receiving post-prostatectomy radiotherapy: 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] Modelling of CTCAE 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+ haematuria. 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 modelling 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 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, 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 two prior studies, with all reported α/β parameter estimates being below 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 five years in the absence of prostate cancer treatment. This may add substantial non-dose-related toxicity noise to the model, even with relatively stringent selection criteria, such as requiring baseline zero GU toxicity.

Late Genitourinary Toxicity in Phase III Hypofractionation Trials

Five major phase III hypofractionation trials have reported cumulative 5 year genitourinary toxicity outcomes, with one announcing 2 year data.[4–6,26–28] These are summarised 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 radiotherapy 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 paper. 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 ranging 2.3 – 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 vs conventional radiotherapy non-inferiority trials, which noted an increase with hypofractionation for late GU G2+ toxicity (relative risk 1.18; 95% CI 0.98 – 1.43; p=0.08), 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 CTCAE late toxicity was worse with ultrahypofractionation (36.25 Gy in 5 fractions) vs conventional or moderately hypofractionated RT (78 Gy in 39 fractions or 62 Gy in 20 fractions) for GU (29% vs 19%), while 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) vs 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 haematuria 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 haematuria 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 modelling 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 (LUTS), 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 radiotherapy [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 LUTS. 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 utilised 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 inter-patient 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 analysing 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 (e.g. HYPO-RT-PC and PACE) would be helpful, to confirm applicability at higher doses per fraction.

Conclusions

We have fitted modified Lyman Kutcher-Burman models for individual genitourinary toxicity endpoints following conventional and moderately hypofractionated external beam radiotherapy. Three models were statistically improved by use of an EQD2 correction: Dysuria G1+, Haematuria G1+ and Haematuria G2+. For these, the fitted α/β ratio was low, ranging from 0.6 – 2.0 Gy, with 95% confidence intervals for these estimates contained within 0.1 – 3.2 Gy. These low estimates suggest that the

therapeutic gain with hypofractionation may be less than expected from modelling using the usual assumption that the late α/β ratio is 3 – 5 Gy. This is important given current trends towards more profoundly hypofractionated radiotherapy. For ultrahypofractionated radiotherapy 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. *et al.* EAU-EANM-ESTRO-ESUR-SIOG Guidelines on Prostate Cancer—2020 Update. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent. *European Urology* **79,** 243–262 (2021).
2. Morgan, S. C. *et al.* Hypofractionated Radiation Therapy for Localized Prostate Cancer: Executive Summary of an ASTRO, ASCO, and AUA Evidence-Based Guideline. *Pract. Radiat. Oncol.* **8,** 354–360 (2018).
3. NICE. Prostate cancer: diagnosis and management. NICE guideline [NG131]. (2019).
4. 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.* **17,** 1047–1060 (2016).
5. Catton, C. N. *et al.* Randomized Trial of a Hypofractionated Radiation Regimen for the Treatment of Localized Prostate Cancer. *J. Clin. Oncol.* **35,** 1884–1890 (2017).
6. Lee, W. R. *et al.* Randomized phase III noninferiority study comparing two radiotherapy fractionation schedules in patients with low-risk prostate cancer. *J. Clin. Oncol.* **34,** 2325–2332 (2016).
7. Fowler, J., Chappell, R. & Ritter, M. Is α/β for prostate tumors really low? *Int. J. Radiat. Oncol. Biol. Phys.* **50,** 1021–1031 (2001).
8. Vogelius, I. R. & Bentzen, S. M. 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.* **100,** 858–865 (2018).
9. Dearnaley, D. & Hall, E. *How will the CHHiP trial affect the future of prostate*

radiotherapy? Expert Review of Anticancer Therapy **18,** 607–609 (Taylor and Francis Ltd, 2018).

10. Brenner, D. J. Fractionation and late rectal toxicity. *Int. J. Radiat. Oncol.* **60,** 1013–1015 (2004).
11. Marzi, S. *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.* **28,** 117 (2009).
12. Tucker, S. L. *et al.* Estimation of α/β for late rectal toxicity based on RTOG 94-06. *Int. J. Radiat. Oncol. Biol. Phys.* **81,** 600–605 (2011).
13. Brand, D. H. *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.* **110,** 596–608 (2021).
14. Dörr, W. & Bentzen, S. M. Late functional response of mouse urinary bladder to fractionated X-irradiation. *Int. J. Radiat. Biol.* **75,** 1307–15 (1999).
15. Bentzen, S. M., Lundbeck, F., Christensen, L. L. & Overgaard, J. Fractionation sensitivity and latency of late radiation injury to the mouse urinary bladder. *Radiother. Oncol.* **25,** 301–307 (1992).
16. Fiorino, C. *et al.* Modelling the impact of fractionation on late urinary toxicity after postprostatectomy radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys.* **90,** 1250–1257 (2014).
17. Cozzarini, C. *et al.* Patient-reported urinary incontinence after radiotherapy for prostate cancer: Quantifying the dose-effect. *Radiother. Oncol.* **125,** 101–106 (2017).
18. Cox, J. D., Stetz, J. & Pajak, T. F. Toxicity criteria of the radiation therapy oncology group (RTOG) and the european organization for research and

treatment of cancer (EORTC). *Pergamon Int. J. Radi ati Oncol. Biol. Phys* **3,** 34–1346 (1995).

19. Dearnaley, D. P. *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.* **83,** 31–41 (2007).
20. LENT SOMA tables. *Radiother. Oncol.* **35,** 1–81 (1995).
21. Tucker, S. L. *et al.* Analysis of Radiation Pneumonitis Risk Using a Generalized Lyman Model. *Int. J. Radiat. Oncol. Biol. Phys.* **72,** 568–574 (2008).
22. Nelder, J. A. & Mead, R. A Simplex Method for Function Minimization. *Comput. J.* **7,** 308–313 (1965).
23. Efron, B. & Tibshirani, R. Improvements on cross-validation: The .632+ bootstrap method. *J. Am. Stat. Assoc.* **92,** 548–560 (1997).
24. Bonferroni, C. E. *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: 8th January 2020)
26. Incrocci, L. *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.* **17,** 1061–1069 (2016).
27. Widmark, A. *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* **394,** 385–395 (2019).

28. Tree, A. C. *et al.* Comparison of side effects at 2 years in the randomised PACE-B trial (SBRT vs standard radiotherapy). in *ESTRO 2021: Presentation Number: OC-0289* (2021).
29. Royce, T. J. *et al.* Conventional Versus Hypofractionated Radiation Therapy for Localized Prostate Cancer: A Meta-analysis of Randomized Noninferiority Trials. *European Urology Focus* **5,** 577–584 (2019).
30. Fransson, P. *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.* **22,** 235–245 (2021).
31. Rammant, E. *et al.* Patient- versus physician-reported outcomes in prostate cancer patients receiving hypofractionated radiotherapy within a randomized controlled trial. *Strahlentherapie und Onkol.* **195,** 393–401 (2019).
32. Marcello, M. *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.* **10,** (2020).
33. Shelley, L. E. A. *et al.* Delivered dose can be a better predictor of rectal toxicity than planned dose in prostate radiotherapy. *Radiother. Oncol.* **123,** 466–471 (2017).

Figures:

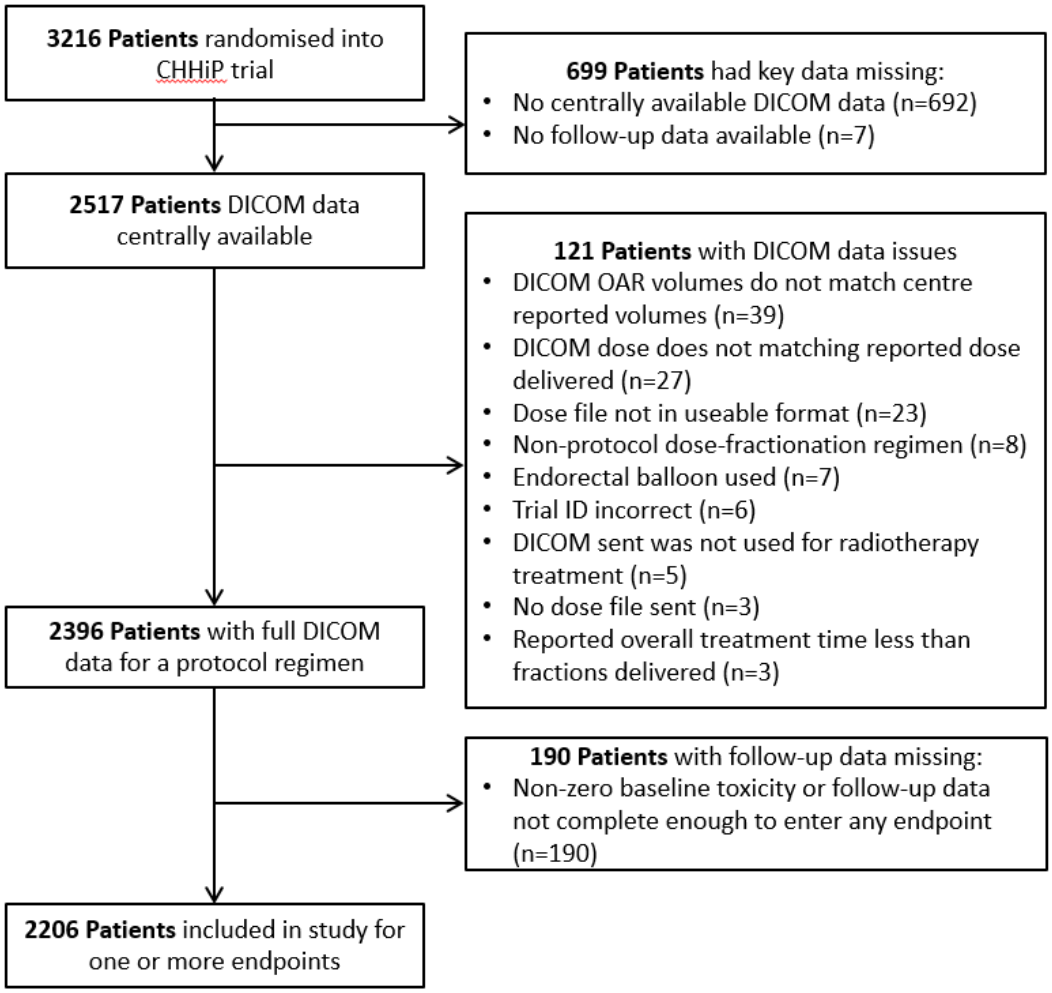

Figure 1. Patient Flow Diagram

Showing any reasons for exclusion of all patients originally randomised into the XXXXXX trial. Abbreviations: DICOM = Digital Imaging and Communications in Medicine; ID = IDentity; OAR = Organ At Risk;

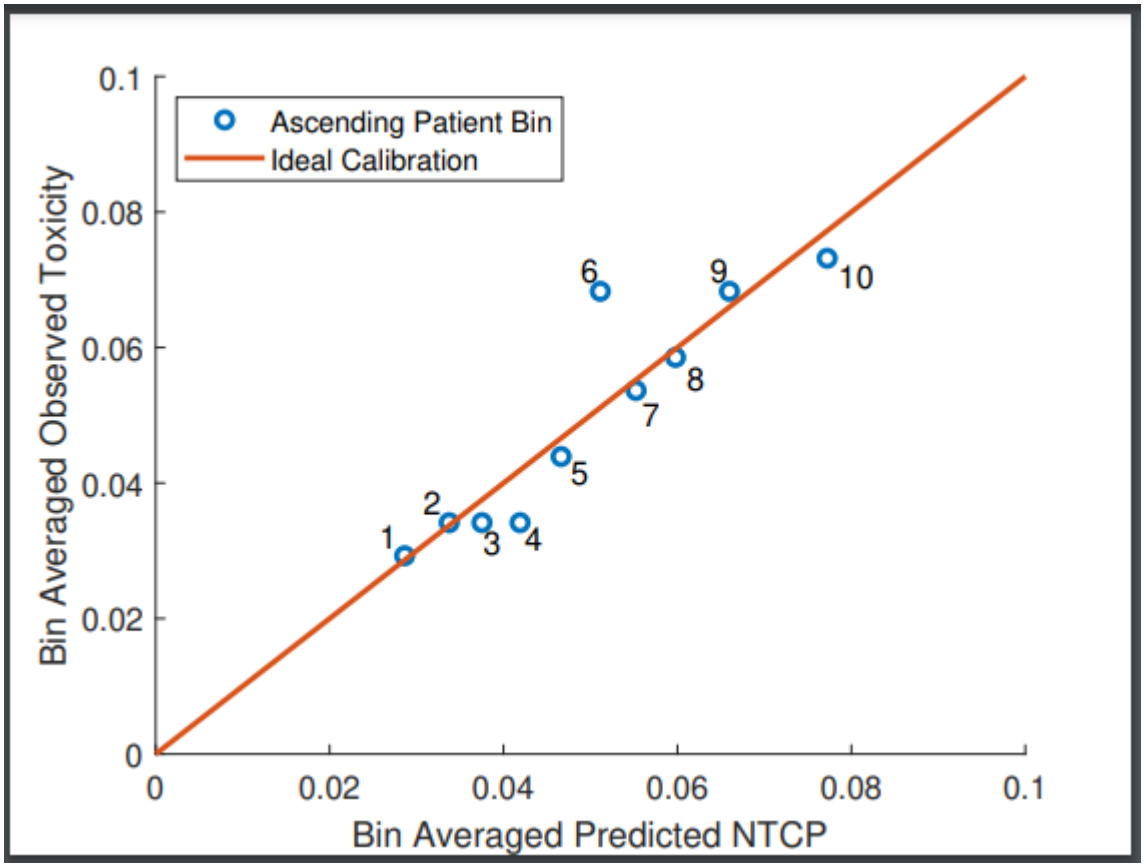

Figure 2. Calibration Plot for LKB-EQD2 Haematuria G2+ Model

Binned decile calibration plot for Haematuria 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 2Gy/fraction Correction; NTCP = Normal Tissue Complication Probability.

Tables

Table 1. Comparison of Baseline Characteristics Between This Substudy and Whole XXXXXX Trial Population

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. Chi-square 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.

Characteristic	**Bladder α/β Ratio Substudy**		**Whole XXXXXX Trial**		**Chi-Square Goodness of Fit**
	Median	**Range**	**Median**	**Range**	
Age	69 years	52-80 years	69 years	44-85 years	
Arm	**No.**	**%**	**No.**	**%**	
57Gy/19f	751	34%	1077	33%	
60Gy/20f	749	34%	1074	33%	0.54
74Gy/37f	706	32%	1065	33%	
NCCN Risk Group					
Low risk	306	14%	484	15%	
Intermediate risk	1,652	75%	2347	73%	0.12
High risk	248	11%	385	12%	
Gleason score					
≤6	747	34%	1122	35%	
7	1,394	63%	1995	62%	0.53
8	65	3%	99	3%	
Clinical T Stage					
T1	849	39%	1170	36%	
T2	1,191	54%	1766	55%	**0.04**
T3	165	8%	277	9%	
Missing	1	0%	3	<1%	
Pre-ADT PSA					
<10 ng/mL	1,075	49%	1567	49%	
10-20 ng/mL	1,005	46%	1415	44%	0.22
≥20 ng/mL	126	6%	208	6%	
Missing	0	0%	26	<1%	
Comorbidities					
Diabetes (yes)	225	10%	342	11%	0.51
Hypertension (yes)	866	39%	1276	40%	0.69
Pelvic Surgery (yes)	165	8%	252	8%	0.53
Prior TURP (yes)	174	8%	259	8%	0.77
Total	2,206	100%	3216	100%	

Table 2. LKB-EQD2 Model Fits for All Endpoints

Fits of the 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 penalising overfitting.

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	**0.0046**
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)	-475.5	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	**0.034**
Haematuria 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	**0.015**
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	0.24
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	0.35
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

Table 3. Acute Toxicity GU G2+ as Dose Modifying Factor

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+, Haematuria G1+, Haematuria 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.

Model	n	m	TD50	α/β ratio	Dose-Modifying Factor	632 Likelihood	Likelihood ratio test p-value
Dysuria G1+ (n=1611)							
LKB-NoEQD2	0.48 (0.01-1.86)	0.76 (0.56-0.85)	253.6 (119.3-1000.0)	0.0 (0.0-0.0)	N/A	-614.5	
LKB-EQD2	0.02 (0.01-0.11)	0.19 (0.12-0.46)	89.9 (81.8-132.6)	1.6 (0.5-2.5)	N/A	-610.4	**0.0044**
LKB-EQD2-DMF	0.02 (0.01-0.78)	0.19 (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**
Haematuria G1+ (n=1576)							
LKB-NoEQD2	1.93 (0.04-10.00)	0.75 (0.67-0.78)	550.4 (166.8-1000)	0.0 (0.0-0.0)	N/A	-524.5	
LKB-EQD2	0.06 (0.01-4.76)	0.41 (0.23-0.73)	141.7 (98.6-346.3)	0.1 (0.1-2.1)	N/A	-523.6	0.16
LKB-EQD2-DMF	0.05 (0.01-0.37)	0.38 (0.23-0.66)	158.4 (106.1-999.8)	0.4 (0.1-262.0)	1.27 (1.09-5.63)	-517.6	**0.00054**
Haematuria G2+ (n=1574)							
LKB-NoEQD2	8.64 (0.01-10.00)	0.59 (0.56-0.59)	1000.0 (274.3-1000)	0.0 (0.0-0.0)	N/A	-311	
LKB-EQD2	0.02 (0.01-0.11)	0.29 (0.18-0.50)	148.9 (106.8-358.3)	0.1 (0.1-1.4)	N/A	-309	**0.048**
LKB-EQD2-DMF	0.01 (0.01-0.10)	0.29 (0.18-0.48)	170.7 (115.6-799.0)	0.1 (0.1-1.3)	1.22 (1.06-2.77)	-305	**0.0046**

Table 4. Summary of Hypofractionation Trials and Genitourinary Toxicity

Table showing the 6 major phase III trials of hypofractionated external beam radiotherapy 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; RTOG = Radiation Therapy Oncology Group; PTV = Planning Target Volume.

Trial [Reference]	Dose- Fractionation Schedule (PTV Doses)		Assumed Late Effects α/β ratio	Toxicity Scale	Cumulative Late Genitourinary Toxicity			
					G2+		G3+	
	Control	HypoFr			Control	HypoFr	Control	HypoFr
XXXXXX **[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