**Article**

# Radiogenomics Consortium Genome-Wide Association Study Meta-analysis of Late Toxicity after Prostate Cancer Radiotherapy

Sarah L. Kerns[1,*], Laura Fachal[2,*], Leila Dorling[3,*], Gillian C. Barnett[3,4], Andrea Baran[5], Derick R. Peterson[5], Michelle Hollenberg[6], Ke Hao[7], Antonio Di Narzo[7], Mehmet Eren Ahsen[7], Gaurav Pandey[7], Søren M. Bentzen[8], Michelle Janelsins[1], Rebecca M. Elliott[9], Paul D.P. Pharoah[4], Neil G. Burnet[9], David P. Dearnaley[10], Sarah L. Gulliford[10], Emma Hall[11], Matthew R. Sydes[12], Miguel E Aguado-Barrera[13], Antonio Gómez-Caamaño[14], Ana M Carballo[14], Paula Peleteiro[14], Ramón Lobato-Busto[15], Richard Stock[16], Nelson N. Stone[17], Harry Ostrer[18], Nawaid Usmani[19], Sandeep Singhal[20], Hiroshi Tsuji[21], Takashi Imai[21], Shiro Saito[22], Rosalind Eeles[23], Kim DeRuyck[24], Matthew Parliament[19], Alison M Dunning[3,**], Ana Vega[13,25,**], Barry S. Rosenstein[26,**], and Catharine M.L. West[9,**] on behalf of the Radiogenomics Consortium

1. Departments of Radiation Oncology and Surgery, University of Rochester Medical Center, Rochester, New York, USA
2. Department of Oncology, Centre for Cancer Genetic Epidemiology, Strangeways Research Laboratory, University of Cambridge, Cambridge, UK

3. Department of Public Health and Primary Care, Centre for Cancer Genetic Epidemiology, Strangeways Research Laboratory, University of Cambridge, Cambridge, UK

4. Department of Oncology, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK

5. Department of Biostatistics and Computational Biology, University of Rochester Medical Center, Rochester, New York, USA

6. Department of Computational Biology, University of Rochester, Rochester, New York, USA

7. Department of Genetics and Genomic Sciences and Icahn Institute for Data Science and Genomic Technology, Icahn School of Medicine at Mount Sinai, New York, New York, USA

8. Division of Biostatistics and Bioinformatics, University of Maryland Greenebaum Cancer Center, and Department of Epidemiology and Public Health, University of Maryland School of Medicine, Baltimore, USA

9. Division of Cancer Sciences, the University of Manchester, Manchester Academic Health Science Centre, Christie Hospital, Manchester, UK

10. Academic UroOncology Unit, The Institute of Cancer Research and Royal Marsden NHS Foundation Trust, London, UK

11. Clinical Trials and Statistics Unit, The Institute of Cancer Research, London, UK

12. MRC Clinical Trials Unit at UCL, Institute of Clinical Trials and Methodology, University College London, London, UK

13. Fundación Pública Galega de Medicina Xenómica-SERGAS & Instituto de Investigación Sanitaria de Santiago de Compostela (IDIS), Santiago de Compostela, Spain

14. Department of Radiation Oncology, Complexo Hospitalario Universitario de Santiago, Servizo Galego de Saúde (SERGAS), Santiago de Compostela, Spain

15. Department of Medical Physics, Complexo Hospitalario Universitario de Santiago, SERGAS, 15706, Santiago de Compostela, Spain

16. Department of Radiation Oncology, Icahn School of Medicine at Mount Sinai, New York, New York, USA

17. Department of Urology, Icahn School of Medicine at Mount Sinai, New York, New York, USA

18. Departments of Pathology and Genetics, Albert Einstein College of Medicine, Bronx, New York, USA

19. Division of Radiation Oncology, Department of Oncology, Cross Cancer Institute, University of Alberta, Edmonton, Canada

20. Department of Pathology and Cell Biology, Columbia University, New York, New York, USA

21. National Institute of Radiological Science, National Institutes for Quantum and Radiological Science and Technology, Chiba, Japan

22. Department of Urology, National Tokyo Medical Center, Tokyo, Japan

23. Division of Genetics and Epidemiology, Institute of Cancer Research and Royal Marsden NHS Foundation Trust, London, UK

24. Departments of Basic Medical Sciences and Radiotherapy, Ghent University Hospital, Ghent, Belgium

25. Grupo de Medicina Xenómica, Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Universidade de Santiago de Compostela, Santiago de Compostela, Spain

26. Departments of Radiation Oncology & Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, New York, USA

*These authors contributed equally to this work.

**Co-senior authors who jointly directed this work.

**Author for correspondence:** Catharine M.L. West, PhD., Institute of Cancer Sciences, the University of Manchester, Manchester Academic Health Science Centre, Christie Hospital, Manchester, M20 4BX, UK. Email: Catharine.West@manchester.ac.uk. Phone: +44 (0) 161 446 8275.

**Abstract**

**Background:** 10-20% of patients develop long-term toxicity following radiotherapy for prostate cancer. Identification of common genetic variants associated with susceptibility to radiotoxicity might improve risk prediction and inform functional mechanistic studies.

**Methods:** We conducted an individual patient data meta-analysis of six genome-wide association studies (n=3,871) in men with European ancestry who underwent radiotherapy for prostate cancer. Radiotoxicities (increased urinary frequency, decreased urinary stream, hematuria, rectal bleeding) were graded prospectively. Grouped relative risk models tested associations with ~6 million genotyped/imputed variants (time to first ≥grade 2 toxicity event). Variants with two-sided $P_{meta}<5\times10^{-8}$ were considered statistically significant. Bayesian false discovery probability provided an additional measure of confidence. Statistically significant variants were evaluated in three Japanese cohorts (n=962). All statistical tests were two-sided.

**Results:** Meta-analysis of the European ancestry cohorts identified three genomic signals: single nucleotide polymorphism (SNP) rs17055178 with rectal bleeding ($P_{meta}=6.2\times10^{-10}$), rs10969913 with decreased urinary stream ($P_{meta}=2.9\times10^{-10}$) and rs11122573 with hematuria ($P_{meta}=1.8\times10^{-8}$). Fine scale mapping of these three regions identified another independent signal (rs147121532) associated with hematuria ($P_{conditional}=4.7\times10^{-6}$). Credible causal variants at these four signals lie in gene-regulatory regions, some modulating expression of nearby genes. Previously identified variants showed consistent associations (rs17599026 with increased urinary frequency, rs7720298 with decreased urinary stream, rs1801516 with overall toxicity) in new

5

cohorts. rs10969913 and rs17599026 had similar effects in the photon-treated

Japanese cohorts.

**Conclusions:** This study increases understanding of the architecture of common

genetic variants affecting radiotoxicity, points to novel radio-pathogenic mechanisms,

and develops risk models for testing in clinical studies. Further multi-national

radiogenomics studies in larger cohorts are worthwhile.

Long-term side-effects following radiotherapy impact the health-related quality-of-life for cancer survivors [1]. Radiation dose and irradiated volume are the most important factors affecting toxicity risk but are broadly similar within patient populations. Known modifiers of the relationship between radiation exposure and risk of radiotoxicity include patient age, smoking history, concurrent treatments and co-morbidities. However, considerable inter-individual variation in radiotoxicity remains unexplained after allowing for dosimetric and patient-level factors. An individual's response to radiation is a heritable trait as evidenced by the similar cellular radiosensitivity of related individuals [2], intra-individual correlation in tissue response to therapeutic radiation [3], and the observation that rare mutations in some genes increase risk of radiotoxicity [4]. Evidence suggests common variants may explain some of the remaining inter-individual variation in radiotoxicity susceptibility [2, 5]. Simulation shows that the accuracy of models to predict radiotoxicity is improved when genetic risk variants are combined with clinical and dosimetric parameters [6].

Genetic predisposition to radiotoxicity in non-syndromic individuals is poorly understood. Pre-clinical studies suggest that the biologic mechanisms are complex involving cell death, premature senescence, inflammation, tissue remodeling with development of fibrosis, and vascular damage. Large genetic studies are difficult because it is challenging to build cohorts associated with a phenotype that takes months to years to develop. In addition, the radiosensitivity phenotype varies by tumor site and means of assessment.

In prostate cancer patients, common radiotoxicities include increased urinary frequency, radiation cystitis (characterized by urinary bleeding, pain, and inflammation),

urinary retention, and rectal bleeding [7, 8]. Radiotoxicity prevalence in prostate cancer survivors varies by radiation delivery modality, time of follow-up, and method of symptom assessment (particularly for clinician vs patient scores). A large (N=1,571) prospective cohort study assessing clinician-assigned toxicity using the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) reported the actuarial likelihood at 10 years of 15% and 9% for ≥ grade 2 urinary and rectal toxicities, respectively [9].

Despite ~50% of cancer patients undergoing radiotherapy, the collection of radiotoxicity data is sporadic and inconsistent. Most centers do not collect data routinely with the detail and standardization required to conduct radiogenomic studies. There is also a clear potential for clinical impact, e.g., with an option for modified radiotherapy planning as increasingly conformal dose delivery techniques become available. The Radiogenomics Consortium was established to facilitate data sharing and develop methods for meta-analyses. Our first pooled genome wide association study (GWAS) performed in prostate cancer survivors demonstrated that it was possible to meta-analyze data and identify genomic risk regions, despite the heterogeneity inherent in radiotherapy cohorts [11].

As part of a long-term goal to identify sufficient variants to develop a polygenic risk model for radiotoxicity, we undertook a larger meta-analysis on six cohorts comprising 3,871 prostate cancer survivors. Secondary aims were to validate previously published SNPs and evaluate risk SNPs in Japanese cohorts. STROGAR guidelines [12] for reporting radiogenomic studies, which build on the STREGA and STROBE guidelines [13, 14], were followed throughout.

8

**Methods**

*Study Participants*

  The study included individuals with prostate adenocarcinoma, treated with radiotherapy with curative intent, and followed prospectively for development of toxicity. All participants gave written informed consent, and cohorts were collected following standards indicated by the Declaration of Helsinki. Participants were from either hospital-based cohorts or clinical trials (**Supplementary Methods**). **Supplementary Figure 1** summarizes the six mainly European-ancestry cohorts (RAPPER, RADIOGEN, GenePARE, UGhent, CCI-BT, CCI-EBRT) included in the meta-analysis and three Japanese cohorts (PRRG-photon, N=170; PRRG-Cion, N=538; and NTMC, N=254; total N=962) analyzed as a separate replication set. Individuals were excluded if: DNA samples/genotyping failed quality control measures; they had non-European (or non-Japanese) ancestry determined using a preselected panel of ancestry informative markers [15]; and/or data were unavailable on androgen deprivation therapy, prior prostatectomy, age at treatment, and total biological effective dose (BED). A previously published GWAS meta-analysis [11] included a subset of RAPPER (RAPPER-I), RADIOGEN, GenePARE (GenePARE-I), and CCI-EBRT. RAPPER-II, GenePARE-II, UGhent, CCI-BT, and the Japanese cohorts were not analyzed previously.

*Assessment of late radiotherapy toxicity*

  Toxicity was assessed from six months up to 5-years after the start of radiotherapy, with the exception of the UGhent cohort (follow-up from 18-30 months

only). Different toxicity assessments were used (**Supplementary Table 1**). Toxicity

grades were harmonized to the CTCAE scale (**Supplementary Table 1**). Assessment

times were coarsened into discrete intervals for time-to-event analysis. Six-month

intervals were used except for RAPPER which used 12-month intervals after the first 2

years. Four toxicity outcomes were analyzed: increased urinary frequency, decreased

urinary stream, hematuria, and rectal bleeding. We also tested a measure of overall

toxicity using STAT score [21].


*Genotyping, Quality Control and Imputation*

Germline DNA from whole blood was genotyped for published GWASs [5, 10, 22-

25] using Affymetrix SNP6.0 arrays (Affymetrix, Inc.; Santa Clara, CA; CCI-EBRT and

GenePARE-I) or Illumina CytoSNP12 arrays (Illumina, Inc.; San Diego, CA; RAPPER-

1). Everything else was genotyped using Illumina OncoArray-500K BeadChips (Illumina,

Inc.; San Diego, CA). Methods used by PRACTICAL were followed [15]. Datasets were

processed and imputed using reference haplotypes from the1000 Genomes Project

(see **Supplementary Methods**).


*Statistical Analysis - GWAS Meta-Analysis*

Genetic variants were tested for associations with toxicity using a grouped

relative risk model adjusting for androgen deprivation therapy, prior prostatectomy, age

at treatment, and total BED. Co-variables were selected *a priori* to reduce

heterogeneity. A per-allele additive genetic model was assumed. In RAPPER [26] and

GenePARE [23, 27], where samples were genotyped in two batches using different

arrays, batch was included as a binary adjustment variable. Outcomes were defined as time from the start of radiation to onset of first occurrence of grade ≥2 toxicity (see **Supplementary Table 1**) with time coarsened into discrete intervals. Efron's method was used to handle ties.

A fixed-effects meta-analysis using inverse variance weighting combined genetic variant-toxicity associations across studies. Variants were considered statistically significant if the two-sided meta-analysis P-value ($P_{meta}$) was $<5 \times 10^{-8}$ and the chi-squared test of heterogeneity p-value was >0.05. The likelihood that an association nominally statistically significant at a given threshold depends not only on the P-value, but also on the power to detect a given association. We therefore assessed the likelihood that our statistically significant associations represent false positives using the Bayes false discovery probability [28]. **Supplementary Table 2** reports power analysis for a range of minor allele frequencies and per-allele effect sizes. All statistical tests were two-sided.

*Fine-scale mapping and credible causal variant (CCV) annotation*

Genomic regions surrounding each statistically significant association were fine-mapped using conditional analysis (details in **Supplementary Materials**). CCVs were annotated with Variant Effect Predictor [31] to determine their effect on genes, transcripts, and protein sequences and overlapped with Encode enhancer-like and promoter-like regions [31-33]. Potential eQTLs were identified using the GTEx Portal.

*Statistical Analysis - Multivariable Modeling*

11

Clinical and genetic variables were combined using cohort-stratified grouped relative risk models. Stepwise model selection was used to identify parsimonious multivariable models for each toxicity outcome. Confidence intervals and p-values were likelihood based and two-sided, with p-values ≤ 0.05 considered statistically significant. The proportional hazards assumption for each predictor was tested at the nominal 2-sided 0.05 level, one at a time, by extending the model to allow separate hazard ratios before and after 1095 days via an interaction of each predictor with I(time > 1095 days); all p-values were > 0.05. Additional details are in **Supplementary Methods**.

**Results**

**Table 1** summarizes the characteristics of the 3,871 participants included in the meta-analysis. The cumulative probability of developing radiotoxicity varied by cohort and outcome (**Figure 1** and **Table 1**), and agreed with prior reports [9, 37-40]. As expected, radiotoxicity was higher in patients receiving brachytherapy [1]. Lower toxicities were observed in UGhent, possibly due to differences in grading and/or shorter follow-up time. Bivariate Cox models provide evidence that the three urinary toxicities were associated with each other (e.g. HR=6.01, 95% CI=4.70 to 7.62 for association of decreased urinary stream with increased urinary frequency) whereas no urinary toxicity was associated with rectal bleeding (e.g. HR=0.96, 95% CI=0.52 to 1.60 for association of decreased urinary stream with rectal bleeding). Meta-analysis Q-Q plots (**Supplementary Figure 2**) showed no genomic inflation, suggesting population structure was adequately controlled using principal components analysis to restrict the

study to European ancestry individuals. Three independent SNPs reached statistical significance with a BFDP<2% (**Table 2** and **Figure 2**): rs17055178 with rectal bleeding ($P_{meta}$=6.2x10$^{-10}$), rs10969913 with decreased urinary stream ($P_{meta}$=2.9x10$^{-10}$) and rs11122573 with hematuria ($P_{meta}$=1.8x10$^{-8}$). Variants with MAF 1-4% and SNPs falling just short of our criteria, including several variants associated with a secondary endpoint of overall toxicity, measured using STAT score [21], are listed in **Supplementary Table 3**.

Fine-scale mapping identified CCVs (**Figure 3, Supplementary Tables 4 and 5**). A second independent signal (tagged by rs147121532, $P_{conditional}$=4.7x10$^{-6}$) at chr1:230337180-231337180 was associated with hematuria. The first signal (tagged by rs11122573, **Supplementary Tables 4 and 5**) includes 47 CCVs (together explaining 93% of the posterior probability of risk). These CCVs lie in active enhancer-/promoter-like gene-regulatory regions (**Supplementary Figure 3A**, **Supplementary Table 6**). Their risk alleles decrease expression of *AGT* (encoding angiotensinogen; ENSG00000135744.7) and *COG2* (encoding conserved oligomeric Golgi complex subunit 2; ENSG00000135775.9) in multiple tissues including arteries (**Supplementary Figure 3B**). The second signal with 10 CCVs (tagged by rs147121532; explaining 54% of the posterior probability; **Figure 3**, **Supplementary Tables 4 and 5**) has risk alleles that decrease expression of *CAPN9* (encoding the intestinal protease, calpain-9; ENSG00000135773.8) and *ARV1* (encoding ARE2 required for viability [ARV1] homolog, fatty acid homeostasis modulator; ENSG00000173409.9). The risk alleles in the second signal were also associated with differential expression of two non-coding RNAs: decreased expression of ncRNA AL512328.1 (ENSG00000244137.1), which

13

overlaps partially with *AGT* and *CAPN9*; and increased expression of ncRNA

LOC101927604 (ENSG00000223393.1). At the chr5:156903410-157903410 region

associated with rectal bleeding, a 15 CCV signal accounts for 98% posterior probability

at the region (**Figure 3**, **Supplementary Tables 4 and 5**). CCVs in this region overlap

active enhancer-like regions in gastrointestinal tissues (**Supplementary Figure 3A;**

**Supplementary Table 6**) but none were statistically significantly associated with

differential gene expression in GTEx (all p>0.05). At the chr9:30366808-31366808

region, associated with decreased urinary stream, a single 15 CCV signal accounts for

99% of the posterior probability (**Figure 3**, **Supplementary Tables 4 and 5**). None were

statistically significantly associated with differential gene expression in the tissues

evaluated by GTEx (all p>0.05).

To explore biological mechanisms underpinning radiotoxicity, we computed gene

and pathway scores from the meta-analysis results (**Supplementary Tables 7 and 8**).

Nine pathways were associated (p<0.05) with multiple toxicities (**Supplementary Table**

**8**), suggesting a common biologic mechanism. For example, the 'ECM [extracellular

matrix] pathway' from the Biocarta database was associated with both increased urinary

frequency and hematuria; 'cytokine signaling in immune system' from the Reactome

database was associated with both decreased urinary stream and hematuria.

SNPs previously associated with radiotoxicity were evaluated for replication in

the new cohorts (**Table 3**). Three SNPs showed a consistent association signal

(rs17599026 with increased urinary frequency, rs7720298 with decreased urinary

stream, and rs1801516 with overall toxicity), thought the effect size was attenuated.

SNPs in *TANC1* [22] showed inconsistent results.

14

The new and previously published variants were evaluated in Japanese ancestry cohorts (**Supplementary Table 9**). In NTMC, the minor alleles of 4 SNPs (rs10969913, rs17599026, rs7720298, rs7582141) were associated with a non-statistically significant increased risk of toxicity, with the effect direction consistent with observations in Europeans. The PRRG photon cohort also showed consistent, though non-statistically significant, association signals for rs10969913 and rs17599026. Associations were not replicated in the PRRG carbon-ion cohort.

Multivariable grouped relative risk models identified independent clinical, dosimetric, and/or co-morbidity factors associated with an increased hazard ratio for each toxicity outcome (**Table 4**). Importantly, SNPs remained independently associated with toxicity and with similar effect sizes in every model. The effect sizes for SNPs were similar to those of established dosimetric or patient-related risk factors. Adjustment for the first 7 principal components derived from analysis of European ancestry samples had minimal impact on model parameters. C-statistics show that while model performance was modest, all four models improved slightly with the addition of SNPs (**Supplementary Table 10**).

## Discussion

Our study identified three new genomic regions associated with radiotoxicity. By performing the first radiogenomics fine mapping study, we identified CCVs in each region and demonstrated two independent signals in the region associated with hematuria. The CCVs in this region were associated with differential expression of local

15

protein coding genes (*AGT*, *COG2*, *CAPN9*, *ARV1*) and non-coding RNAs

(AL512328.1, LOC101927604), pointing towards possible functional mechanisms. For

example, *AGT* encoding angiotensinogen is converted to the active enzyme angiotensin

II by angiotensin converting enzyme (ACE). Prior studies suggest angiotensin signaling

may influence radiation-induced blood vessel wall injury and interstitial fibrosis [41], and

animal and human studies suggest that ACE inhibitors could be radio-protective [42-46].

The regions associated with rectal bleeding and decreased urinary stream did not

contain CCVs associated with differential expression of nearby genes in the available

tissue databases, and may associate with long-distance gene regulation. The rectal

bleeding locus overlaps with active enhancer-like regions in gastrointestinal tissues, and

further studies are needed to uncover the functional effects.

Pathway analysis might also identify new mechanisms involved in the

pathogenesis of radiotoxicity. For example, the top-ranking pathway associated with

hematuria was platelet adhesion to exposed collagen. Platelet adhesion is the first step

in the formation of a platelet plug during hemostasis in response to blood vessel wall

injury [47]. Collagens are involved in the process and some are abundant in vascular

epithelia [48]. Several collagen-binding proteins are expressed on platelets including

integrins [47]. Interestingly, the integrin pathway ranked eighth for an association with

rectal bleeding.

Our analysis is the largest GWAS of late radiotoxicity. While heterogeneity

across cohorts adds noise and masks statistically significant SNP-toxicity associations,

our analysis plan addressed heterogeneity. Use of individual patient data meta-analysis

maximizes statistical power when combining information across multiple studies [49,

50]. In addition, we stratified our multivariable models by study to account for differences in toxicity incidence and other unmeasured covariates that may differ across study populations. Analysis of previously identified risk SNPs provides support for the original discovery as the odds ratios, the best estimate of association, are in the same direction in the new studies. As is commonly observed, the initial effect size estimates for the prior GWAS-identified SNPs were upwardly biased (so-called "winner's curse"), and evaluation here enabled estimation of effect sizes that more likely reflect their true contribution to risk of toxicity. Although unable to replicate the association at 2q24.1 within *TANC1*, it is challenging due to the rarity of minor alleles within Europeans [30]. However, ongoing laboratory studies support a role of *TANC1* in radiation response (personal communication from A.V.), highlighting the importance of functional studies as complementary to association studies.

The multivariable risk models demonstrate that genetic variants, treatment variables, and other clinical factors are independent predictors for radiotoxicity, supporting the suggestion that common variants can improve traditional normal tissue complication probability models [51, 52]. While the c-statistics for our models are modest, SNPs improved model performance. As cohort and sample sizes increase, further GWAS should uncover additional risk SNPs, and sequencing studies will uncover rare, possibly high-penetrance, variants. We foresee the eventual incorporation of polygenic scores or gene panel information into integrative models of radiotoxicity that also include clinical risk factors. While important to continue efforts to identify additional risk SNPs and rare variants, our models are ready for validation, a critical next step towards clinical testing. The REQUITE prospective cohort study [53] of late

radiotoxicity will provide an excellent opportunity to test these risk models. In addition, investigators are actively developing novel clinical trial approaches for testing the ability of risk models to personalize treatment and improve outcomes [54]. We foresee these models being used, e.g., to identify the small proportion of individuals with the highest risk for toxicity who might choose to avoid radiotherapy if watchful waiting or surgery are alternatives; for dose-escalation in patients with a low-risk of toxicity; and to evaluate how model-directed modification of planned doses to normal tissues impacts risk for developing toxicity.

A strength of our study is the prospective longitudinal assessment of toxicity enabling use of time-to-event analysis to maximize information across multiple toxicity assessments. Long-term follow-up is clearly important for radiogenomic studies, and future work should use longitudinal analysis when possible. A second strength of our work is the inclusion of Japanese ancestry cohorts. Most GWAS to date focused on populations of European ancestry, as ethnicity inflates type I error rates and reduces statistical power due to population heterogeneity in allelic effects on a trait [55]. It is important to understand how knowledge gained from European populations transfers to other ethnicities. Methods are being developed to detect genetic variants associated with complex traits allowing for population heterogeneity [56]. Trans-ethnic studies suggest susceptibility loci for traits are generally shared between European and East Asians [57] and, because of the larger sample size, cross-population meta-analyses increase statistical power to detect novel loci [58]. Our cohort sizes are still too small to identify heterogeneity in allelic effects between ethnic groups. However, we performed

the first analysis exploring transferability of SNP-radiotherapy toxicity associations across ethnicities.

Limitations include the lack of detailed dosimetry and comorbidity data for all cohorts, although our stratified analysis maximized use of data across all studies for multivariable models. In addition, while this is the largest GWAS of its kind to date, the sample size is modest compared with disease susceptibility GWAS, and was only powered to detect SNP-toxicity associations with relatively large effect sizes. Our study was too small for multi-SNP modeling, such as polygenic risk score and machine learning-based methods. The latter were successful for other polygenic traits and diseases [59] and should be re-evaluated as larger radiotoxicity cohorts become available. In addition, this study was not designed to identify rare variants associated with radiotherapy toxicity, which requires next generation sequencing of large cohorts or case-control studies.

In summary, by performing the largest GWAS meta-analysis and first fine-mapping study in radiogenomics we identified four new regions associated with radiotoxicity in prostate cancer. We showed the signals affect gene regulation rather than gene coding sequences. This study increases understanding of the architecture of common genetic variants affecting radiotoxicity, and demonstrates that further multi-national radiogenomics studies in larger cohorts are worthwhile.

20

Hospital and Institute of Cancer Research Biomedical Research Centre. CMLW was supported by the NIHR Manchester Biomedical Research Centre.

**Notes**

**References**

1.	Sanda MG, Dunn RL, Michalski J*, et al.* Quality of life and satisfaction with outcome among prostate-cancer survivors. N Engl J Med 2008;358(12):1250-61.

2.	Wu X, Spitz MR, Amos CI*, et al.* Mutagen sensitivity has high heritability: evidence from a twin study. Cancer Res 2006;66(12):5993-6.

3.      Safwat A, Bentzen SM, Turesson I*, et al.* Deterministic rather than stochastic factors explain most of the variation in the expression of skin telangiectasia after radiotherapy. Int J Radiat Oncol Biol Phys 2002;52(1):198-204.

4.      Barnett GC, West CM, Dunning AM*, et al.* Normal tissue reactions to radiotherapy: towards tailoring treatment dose by genotype. Nat Rev Cancer 2009;9(2):134-42.

5.      Barnett GC, Thompson D, Fachal L*, et al.* A genome wide association study (GWAS) providing evidence of an association between common genetic variants and late radiotherapy toxicity. Radiother Oncol 2014;111(2):178-85.

6.      Kerns SL, Kundu S, Oh JH*, et al.* The Prediction of Radiotherapy Toxicity Using Single Nucleotide Polymorphism-Based Models: A Step Toward Prevention. Semin Radiat Oncol 2015;25(4):281-91.

7.      Hauer-Jensen M, Denham JW, Andreyev HJ. Radiation enteropathy-- pathogenesis, treatment and prevention. Nat Rev Gastroenterol Hepatol 2014;11(8):470-9.

8.      Marks LB, Carroll PR, Dugan TC*, et al.* The response of the urinary bladder, urethra, and ureter to radiation and chemotherapy. Int J Radiat Oncol Biol Phys 1995;31(5):1257-80.

9.      Zelefsky MJ, Levin EJ, Hunt M*, et al.* Incidence of late rectal and urinary toxicities after three-dimensional conformal radiotherapy and intensity-modulated radiotherapy for localized prostate cancer. Int J Radiat Oncol Biol Phys 2008;70(4):1124-9.

10. Schumacher FR, Al Olama AA, Berndt SI, *et al.* Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci. Nat Genet 2018; 10.1038/s41588-018-0142-8.

11. Kerns SL, Dorling L, Fachal L, *et al.* Meta-analysis of Genome Wide Association Studies Identifies Genetic Markers of Late Toxicity Following Radiotherapy for Prostate Cancer. EBioMedicine 2016;10:150-63.

12. Kerns SL, de Ruysscher D, Andreassen CN, *et al.* STROGAR - STrengthening the Reporting Of Genetic Association studies in Radiogenomics. Radiother Oncol 2014;110(1):182-8.

13. Little J, Higgins JP, Ioannidis JP, *et al.* Strengthening the reporting of genetic association studies (STREGA): an extension of the STROBE Statement. Hum Genet 2009;125(2):131-51.

14. von Elm E, Altman DG, Egger M, *et al.* The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann Intern Med 2007;147(8):573-7.

15. Amos CI, Dennis J, Wang Z, *et al.* The OncoArray Consortium: A Network for Understanding the Genetic Architecture of Common Cancers. Cancer Epidemiol Biomarkers Prev 2017;26(1):126-135.

16. LENT SOMA scales for all anatomic sites. Int J Radiat Oncol Biol Phys 1995;31(5):1049-91.

17. Dearnaley DP, Hall E, Lawrence D, *et al.* Phase III pilot study of dose escalation using conformal radiotherapy in prostate cancer: PSA control and side effects. Br J Cancer 2005;92(3):488-98.

18.     *RTOG/EORTC Late Radiation Morbidity Scoring Schema.*

https://www.rtog.org/ResearchAssociates/AdverseEventReporting/RTOGEORTCLateR

adiationMorbidityScoringSchema.aspx.

19.     National Cancer Institute. Common Terminology Criteria for Adverse Events

v4.0. In. *NIH publication # 09-7473*; NCI, NIH, DHHS. May 29, 2009.

20.     Barry MJ, Fowler FJ, Jr., O'Leary MP*, et al.* The American Urological Association

symptom index for benign prostatic hyperplasia. The Measurement Committee of the

American Urological Association. J Urol 1992;148(5):1549-57; discussion 1564.

21.     Barnett GC, West CM, Coles CE*, et al.* Standardized Total Average Toxicity

score: a scale- and grade-independent measure of late radiotherapy toxicity to facilitate

pooling of data from different studies. Int J Radiat Oncol Biol Phys 2012;82(3):1065-74.

22.     Fachal L, Gomez-Caamano A, Barnett GC*, et al.* A three-stage genome-wide

association study identifies a susceptibility locus for late radiotherapy toxicity at 2q24.1.

Nat Genet 2014;46(8):891-4.

23.     Kerns SL, Stock R, Stone N*, et al.* A 2-stage genome-wide association study to

identify single nucleotide polymorphisms associated with development of erectile

dysfunction following radiation therapy for prostate cancer. Int J Radiat Oncol Biol Phys

2013;85(1):e21-8.

24.     Kerns SL, Stock RG, Stone NN*, et al.* Genome-wide association study identifies

a region on chromosome 11q14.3 associated with late rectal bleeding following

radiation therapy for prostate cancer. Radiother Oncol 2013;107(3):372-6.

25.    Kerns SL, Stone NN, Stock RG, *et al.* A 2-stage genome-wide association study to identify single nucleotide polymorphisms associated with development of urinary symptoms after radiotherapy for prostate cancer. J Urol 2013;190(1):102-8.

26.    Burnet NG, Elliott RM, Dunning A, *et al.* Radiosensitivity, radiogenomics and RAPPER. Clin Oncol (R Coll Radiol) 2006;18(7):525-8.

27.    Ho AY, Atencio DP, Peters S, *et al.* Genetic predictors of adverse radiotherapy effects: the Gene-PARE project. Int J Radiat Oncol Biol Phys 2006;65(3):646-55.

28.    Wakefield J. A Bayesian measure of the probability of false discovery in genetic epidemiology studies. Am J Hum Genet 2007;81(2):208-27.

29.    Marchini J, Howie B, Myers S, *et al.* A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet 2007;39(7):906-13.

30.    Genomes Project C, Auton A, Brooks LD, *et al.* A global reference for human genetic variation. Nature 2015;526(7571):68-74.

31.    McLaren W, Gil L, Hunt SE, *et al.* The Ensembl Variant Effect Predictor. Genome Biol 2016;17(1):122.

32.    Consortium EP. An integrated encyclopedia of DNA elements in the human genome. Nature 2012;489(7414):57-74.

33.    Sloan CA, Chan ET, Davidson JM, *et al.* ENCODE data at the ENCODE portal. Nucleic Acids Res 2016;44(D1):D726-32.

34.    Pruim RJ, Welch RP, Sanna S, *et al.* LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics 2010;26(18):2336-7.

35.    Aulchenko YS, Struchalin MV, van Duijn CM. ProbABEL package for genome-wide association analysis of imputed data. BMC Bioinformatics 2010;11:134.

36.     Lamparter D, Marbach D, Rueedi R*, et al.* Fast and Rigorous Computation of Gene and Pathway Scores from SNP-Based Summary Statistics. PLoS Comput Biol 2016;12(1):e1004714.

37.     Hamdy FC, Donovan JL, Lane JA*, et al.* 10-Year Outcomes after Monitoring, Surgery, or Radiotherapy for Localized Prostate Cancer. N Engl J Med 2016;375(15):1415-1424.

38.     Odrazka K, Dolezel M, Vanasek J*, et al.* Late toxicity after conformal and intensity-modulated radiation therapy for prostate cancer: impact of previous surgery for benign prostatic hyperplasia. Int J Urol 2010;17(9):784-90.

39.     Resnick MJ, Koyama T, Fan KH*, et al.* Long-term functional outcomes after treatment for localized prostate cancer. N Engl J Med 2013;368(5):436-45.

40.     Syndikus I, Morgan RC, Sydes MR*, et al.* Late gastrointestinal toxicity after dose-escalated conformal radiotherapy for early prostate cancer: results from the UK Medical Research Council RT01 trial (ISRCTN47772397). Int J Radiat Oncol Biol Phys 2010;77(3):773-83.

41.     Rosenbloom J, Castro SV, Jimenez SA. Narrative review: fibrotic diseases: cellular and molecular mechanisms and novel therapies. Ann Intern Med 2010;152(3):159-66.

42.     Moulder JE, Fish BL, Cohen EP. Noncontinuous use of angiotensin converting enzyme inhibitors in the treatment of experimental bone marrow transplant nephropathy. Bone Marrow Transplant 1997;19(7):729-35.

43. Molteni A, Moulder JE, Cohen EF, *et al.* Control of radiation-induced pneumopathy and lung fibrosis by angiotensin-converting enzyme inhibitors and an angiotensin II type 1 receptor blocker. Int J Radiat Biol 2000;76(4):523-32.

44. Charrier S, Michaud A, Badaoui S, *et al.* Inhibition of angiotensin I-converting enzyme induces radioprotection by preserving murine hematopoietic short-term reconstituting cells. Blood 2004;104(4):978-85.

45. Kharofa J, Cohen EP, Tomic R, *et al.* Decreased risk of radiation pneumonitis with incidental concurrent use of angiotensin-converting enzyme inhibitors and thoracic radiation therapy. Int J Radiat Oncol Biol Phys 2012;84(1):238-43.

46. Wedlake LJ, Silia F, Benton B, *et al.* Evaluating the efficacy of statins and ACE-inhibitors in reducing gastrointestinal toxicity in patients receiving radiotherapy for pelvic malignancies. Eur J Cancer 2012;48(14):2117-24.

47. Nuyttens BP, Thijs T, Deckmyn H, *et al.* Platelet adhesion to collagen. Thromb Res 2011;127 Suppl 2:S26-9.

48. Pugh N, Simpson AM, Smethurst PA, *et al.* Synergism between platelet collagen receptors defined using receptor-specific collagen-mimetic peptide substrata in flowing blood. Blood 2010;115(24):5069-79.

49. Cohn LD, Becker BJ. How meta-analysis increases statistical power. Psychol Methods 2003;8(3):243-53.

50. Evangelou E, Ioannidis JP. Meta-analysis methods for genome-wide association studies and beyond. Nat Rev Genet 2013;14(6):379-89.

51.    Tucker SL, Geara FB, Peters LJ*, et al.* How much could the radiotherapy dose be

altered for individual patients based on a predictive assay of normal-tissue

radiosensitivity? Radiother Oncol 1996;38(2):103-13.

52.    El Naqa I, Pandey G, Aerts H*, et al.* Radiation Therapy Outcomes Models in the

Era of Radiomics and Radiogenomics: Uncertainties and Validation. Int J Radiat Oncol

Biol Phys 2018;102(4):1070-1073.

53.    West C, Azria D, Chang-Claude J*, et al.* The REQUITE project: validating

predictive models and biomarkers of radiotherapy toxicity to reduce side-effects and

improve quality of life in cancer survivors. Clin Oncol (R Coll Radiol) 2014;26(12):739-

42.

54.    Azria D, Lapierre A, Gourgou S*, et al.* Data-Based Radiation Oncology: Design of

Clinical Trials in the Toxicity Biomarkers Era. Front Oncol 2017;7:83.

55.    Rosenberg NA, Huang L, Jewett EM*, et al.* Genome-wide association studies in

diverse populations. Nat Rev Genet 2010;11(5):356-66.

56.    Cook JP, Morris AP. Multi-ethnic genome-wide association study identifies novel

locus for type 2 diabetes susceptibility. Eur J Hum Genet 2016;24(8):1175-80.

57.    Akiyama M, Okada Y, Kanai M*, et al.* Genome-wide association study identifies

112 new loci for body mass index in the Japanese population. Nat Genet

2017;49(10):1458-1467.

58.    Li YR, Keating BJ. Trans-ethnic genome-wide association studies: advantages

and challenges of mapping in diverse populations. Genome Med 2014;6(10):91.

59.    Pandey G, Pandey OP, Rogers AJ*, et al.* A Nasal Brush-based Classifier of Asthma Identified by Machine Learning Analysis of Nasal RNA Sequence Data. Sci Rep 2018;8(1):8826.

**Tables**

**Table 1.** Patient characteristics by cohort for the 3,871 individuals included in the GWAS meta-analysis.

| Characteristics | All Cohorts N=3,871 | RAPPER, N=2,010 | RADIOGEN, N=658 | GenePARE, N=492 | UGhent, N=311 | CCI-BT, N=252 | CCI-EBRT, N=148 |
|---|---|---|---|---|---|---|---|
| Age at treatment, median (range)* | 68 (43, 86) | 68 (48, 84) | 72 (47, 86) | 65 (43, 85) | 65 (49, 81) | 65 (45, 79) | 68 (45, 82) |
| NCCN risk group, n (%) | | | | | | | |
|   Very low | 545 (14.1) | 133 (6.6) | 100 (15.2) | 172 (35.0) | 43 (13.8)[†] | 89 (35.3) | 8 (5.4) |
|   Low | 258 (6.7) | 82 (4.1) | 23 (3.5) | 61 (12.4) | 21 (6.8) | 68 (27.0) | 3 (2.0) |
|   Intermediate | 2,635 (68.1) | 1,566 (77.9) | 447 (67.9) | 232 (47.2) | 173 (55.6) | 95 (37.7) | 122 (82.4) |
|   High or Very high | 410 (10.6) | 229 (11.4) | 82 (12.5) | 27 (5.5) | 57 (18.4) | 0 | 15 (10.1) |
|   Not available | 23 (0.6) | 0 | 6 (0.9) | 0 | 17 (5.5) | 0 | 0 |
| Stage at diagnosis, n (%) | | | | | | | |
|   T1a-c, T1x | 1,443 (37.3) | 709 (35.3) | 226 (34.3) | 249 (50.6) | 101 (32.5) | 119 (47.2) | 38 (25.7) |
|   T2a-c, T2x | 2,020 (52.2) | 1,084 (53.9) | 362 (55.0) | 227 (46.1) | 126 (40.5) | 132 (52.4) | 89 (60.1) |
|   T3a-c, T3x | 305 (7.9) | 182 (9.1) | 54 (2.7) | 16 (3.3) | 37 (11.9) | 0 | 16 (10.8) |
|   T4 | 14 (0.4) | 0 | 7 (1.1) | 0 | 6 (1.9) | 0 | 1 (0.7) |
|   Not available | 89 (2.3) | 35 (1.7) | 9 (1.4) | 0 | 41 (13.2) | 1 (0.4) | 4 (2.7) |
| Gleason at diagnosis, n (%) | | | | | | | |
|   ≤6 | 1,702 (44.0) | 605 (30.1) | 403 (61.2) | 310 (63.0) | 142 (45.7) | 212 (84.1) | 30 (20.3) |
|   7 | 1,653 (42.7) | 1,109 (55.2) | 176 (26.8) | 124 (25.2) | 107 (34.4) | 40 (15.9) | 97 (65.5) |
|   ≥8 | 265 (6.8) | 56 (2.8) | 70 (10.6) | 58 (11.8) | 60 (19.3) | 0 | 21 (14.2) |
|   Not available | 251 (6.5) | 240 (11.9) | 9 (1.4) | 0 | 2 (0.6) | 0 | 0 |
| Pre-treatment PSA, median (range) | 8.9 (0, 236.0) | 10.1 (0.6, 33.5) | 9.7 (0.6, 236.0) | 6.2 (0.6, 124.0) | 6.6 (0, 150.0)[‡] | 6.3 (0.5, 16.0) | 10.9 (1.4, 80.0) |
| Radical prostatectomy, n (%)* | | | | | | | |
|   Yes | 225 (5.8) | 0 | 128 (29.5) | 0 | 97 (31.2) | 0 | 0 |
|   No | 3,646 (94.2) | 2,010 (100) | 530 (80.5) | 492 (100) | 214 (68.8) | 252 (100) | 148 (100) |
| Androgen deprivation therapy, n (%)* | | | | | | | |
|   Yes | 3,047 (78.7) | 2,010 (100) | 463 (70.4) | 248 (50.4) | 198 (63.7) | 55 (21.8) | 73 (49.3) |
|   No | 824 (21.3) | 0 | 195 (29.6) | 244 (49.6) | 113 (36.3) | 197 (78.2) | 75 (50.7) |
| Type of radiotherapy, n (%) | | | | | | | |
|   3D-CRT | 895 (25.4) | 237 (11.8) | 658 (100) | 0 | 0 | 0 | 0 |
|   IMRT | 2,239 (57.8) | 1,773 (88.2) | 0 | 7 (1.4) | 311 (100) | 0 | 148 (100) |
|   Brachytherapy | 534 (13.8) | 0 | 0 | 282 (57.3) | 0 | 252 (100) | 0 |
|   Brachytherapy + EBRT | 203 (5.2) | 0 | 0 | 203 (41.3) | 0 | 0 | 0 |
| Total BED[§], median (range)* | 123 (52, 292) | 120 (107, 123) | 123 (57, 127) | 192 (52, 269) | 136 (124, 136) | 158 (80, 292) | 121 (112, 134) |
| Number (%) with grade 2 or worse toxicity | | | | | | | |
|   Increased urinary frequency[‖] | 436 (11.5) | 219 (10.9) | 60 (9.1) | 113 (24.6) | 8 (2.6)[¶] | NA[#] | 15 (10.1) |
|   Decreased urinary stream[**] | 345 (9.9) | 159 (7.9) | 27 (4.1) | 125 (27.2) | NA[††] | NA[#] | 12 (8.1) |
|   Hematuria[‡‡] | 333 (9.2) | 182 (9.1) | 66 (10.0) | 62 (12.6) | 17 (5.5)[¶] | NA[§§] | 6 (4.1) |
|   Rectal bleeding[‖‖] | 423 (12.5) | 273 (13.6) | 79 (12.0) | NA[¶¶] | 6 (1.9)[¶] | 40 (15.9) | 25 (16.9) |

30

\* Age at treatment, radical prostatectomy, androgen deprivation therapy, and total BED were included as covariates in the GWAS meta-analysis. Abbreviations: NCCN, National Comprehensive Cancer Network; PSA, prostate specific antigen; 3D-CRT, three-dimensional conformal radiotherapy; IMRT, intensity modulated radiotherapy; EBRT, external beam radiotherapy (either 3D-CRT or IMRT); BED, biologic effective dose.

† NCCN risk group in the UGhent cohort was defined using pre-radiotherapy PSA rather than PSA at diagnosis.

‡ PSA measurement is pre-radiotherapy but post-prostatectomy in patients who received prior prostatectomy.

§ Total BED was calculated using an α/β ratio of 3 following Stock et al. 2006 IJROBP (Stock et al., 2006).

‖ Increased urinary frequency was evaluable in 3,782 participants with available baseline and follow-up data (2,010 in RAPPER, 658 in RADIOGEN, 459 in GenePARE, 303 in UGhent, and 148 in CCI-EBRT)

¶ Follow-up in UGhent was from 18 months to 30 months as opposed to 6 months to 5 years in all other studies.

# Increased urinary frequency and decreased urinary stream were not analyzed in CCI-BT because assessments were not done at regular intervals.

** Decreased urinary stream was evaluable in 3,470 participants with available baseline and follow-up data (2,010 in RAPPER, 658 in RADIOGEN, 459 in GenePARE, and 148 in CCI-EBRT)

†† Decreased urinary stream was not assessed in UGhent.

‡‡ Hematuria was evaluable in 3,619 participants with available baseline and follow-up data (2,010 in RAPPER, 658 in RADIOGEN, 492 in GenePARE, 311 in UGhent, and 148 in CCI-EBRT)

§§ Hematuria was not assessed in CCI-BT.

‖‖ Rectal bleeding was evaluable in 3,379 participants with available baseline and follow-up data (2,010 in RAPPER, 658 in RADIOGEN, 311 in UGhent, 252 in CCI-BT, and 148 in CCI-EBRT)

¶¶ Rectal bleeding was assigned a single grade in GenePARE using information across all follow up assessments, and so this outcome was not available for analysis

**Table 2**. Study-specific and overall results for new risk SNPs identified via GWAS meta-analysis[*] of six European ancestry cohorts. Bold values correspond to results from meta-analysis.

| Genetic variant | Chr[†] | Minor Allele | MAF [‡] | Toxicity outcome | Study | Info[§] | Mean minor allele dosage | | HR (95% CI) [ǁ] | $P_{meta}$[¶] | $P_{het}$[#] | BFDP [**], % |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Toxicity | No Toxicity | | | | |
| rs17055178 | chr5:157,403,410 | G | 0.09 | Time to first grade 2+ rectal bleeding | Meta-analysis | - | | | **1.95 (1.58 to 2.40)** | **6.2x10⁻¹⁰** | **0.61** | **0.09** |
| | | | | | RAPPER | 0.81, 0.99 | 0.22 | 0.13 | 1.78 (1.37 to 2.32) | | | |
| | | | | | RADIOGEN | 0.99 | 0.33 | 0.14 | 2.58 (1.69 to 3.95) | | | |
| | | | | | GenePARE | NA[††] | NA[††] | NA[††] | NA[††] | | | |
| | | | | | UGhent[‡‡] | 0.99 | 0.17 | 0.14 | 1.38 (0.18 to 10.4) | | | |
| | | | | | CCI-BT | 0.99 | 0.25 | 0.14 | 2.01 (0.97 to 4.20) | | | |
| | | | | | CCI-EBRT | 0.98 | 0.12 | 0.13 | 1.27 (0.38 to 4.25) | | | |
| rs10969913 | chr9:30,866,808 | G | 0.05 | Time to first grade 2+ decreased urinary stream | Meta-analysis | - | | | **3.92 (2.57 to 6.00)** | **2.9x10⁻¹⁰** | **0.08** | **1.07** |
| | | | | | RAPPER | 0.61, 0.95 | 0.04 | 0.02 | 1.86 (0.76 to 4.54) | | | |
| | | | | | RADIOGEN | 0.95 | 0.04 | 0.02 | 2.03 (0.27 to 15.4) | | | |
| | | | | | GenePARE | 0.99, 0.95 | 0.11 | 0.04 | 4.36 (2.55 to 7.46) | | | |
| | | | | | UGhent | NA[§§] | NA[§§] | NA[§§] | NA[§§] | | | |
| | | | | | CCI-BT | NA[ǁǁ] | NA[ǁǁ] | NA[ǁǁ] | NA[ǁǁ] | | | |
| | | | | | CCI-EBRT | 0.95 | 0.27 | 0.02 | 14.3 (3.78 to 54.4) | | | |
| rs11122573 | chr1:230,837,180 | T | 0.06 | Time to first grade 2+ hematuria | Meta-analysis | - | | | **1.92 (1.53 to 2.42)** | **1.8x10⁻⁸** | **0.14** | **1.96** |
| | | | | | RAPPER | 0.99 | 0.19 | 0.14 | 1.42 (0.99 to 2.04) | | | |
| | | | | | RADIOGEN | 0.99 | 0.34 | 0.15 | 2.40 (1.54 to 3.73) | | | |
| | | | | | GenePARE | 0.99, 0.99 | 0.18 | 0.11 | 2.01 (1.25 to 3.22) | | | |
| | | | | | UGhent | 0.99 | 0.47 | 0.14 | 3.59 (1.72 to 7.49) | | | |
| | | | | | CCI-BT | NA[¶¶] | NA[¶¶] | NA[¶¶] | NA[¶¶] | | | |
| | | | | | CCI-EBRT | 1.000 | 0.17 | 0.16 | 0.99 (0.13 to 7.58) | | | |

[*] Within each cohort, SNP-toxicity associations were adjusted for age at treatment, prior prostatectomy, adjuvant hormonal therapy, and total BED. Associations in RAPPER and GenePARE were also adjusted for genotyping batch. Abbreviations: SNP, single nucleotide polymorphism; HR, hazard ratio; CI, confidence interval; BFDP, Bayesian false discovery probability; MAF, minor allele frequency; NA, not analyzed.

[†] Base position is according to Genome Reference Consortium Human Build 37 (hg19).

[‡] Minor allele frequency for each is from PRACTICAL Oncoarray samples of European ancestry

[§] Imputation info score values in RAPPER are from the cytoSNP12 array and oncoarray respectively; values in Gene-PARE are from the AffySNP6.0 array and oncoarray respectively; values in all other studies are from the oncoarray.

[ǁ] Hazard ratio corresponds to the minor allele with the major allele treated as the reference group.

[¶] Two-sided $P_{meta}$ was calculated using a Wald test.

[#] Two-sided heterogeneity p-value was calculated using a Chi-square test.

[**] BFPD estimated assuming a prior variance, W = 0.32^2, and prior probability of a non-null association 0.0001.

[††] Rectal bleeding was assigned a single grade in GenePARE using information across all follow up assessments, and so this outcome was not available for analysis using time-to-event analysis.

[‡‡] There were only 6 rectal bleeding events in UGhent. Exclusion of this cohort from meta-analysis had minimal impact on the results: HR 1.95, 95% CI 1.58 to 2.41, $P_{meta}$ 6.1x10⁻¹⁰.

[§§] Decreased urinary stream was not assessed in UGhent.

[ǁǁ] Increased urinary frequency and decreased urinary stream were not analyzed in CCI-BT because assessments were not done at regular intervals.

32

¶¶ Hematuria was not assessed in CCI-BT.

**Table 3**. Association results for risk loci identified in prior genetic association studies.

| Genetic variant and Gene symbol | Chr[†] | Minor allele | MAF[‡] | Toxicity outcome | Results from prior publication | | | Meta-analysis of new studies not included in prior publication | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | OR (95% CI) | $P_{meta}$[§] | N | Study, N | Info[*] | OR (95% CI) | $P_{meta}$[§] |
| rs17599026 *KDM3B* | chr5:137,763,798 | T | 0.07 | Presence of grade 1+ increased urinary frequency at 2 years after radiotherapy | 3.12 (2.08-4.69) | $4.2 \times 10^{-8}$ | 1,564 | Meta-analysis<br>RAPPER-II, N=1,255<br>GenePARE-II, N=161<br>UGhent, N=281 | -<br>0.96<br>0.96<br>0.96 | 1.23 (0.91 to 1.67)<br>1.27 (0.90 to 1.80)<br>1.10 (0.45 to 2.69)<br>1.08 (0.44 to 2.64) | 0.19 |
| rs7720298 *DNAH5* | chr5:13,858,328 | G | 0.30 | Presence of grade 1+ decreased urine stream at 2 years after radiotherapy | 2.71 (1.90-3.86) | $3.2 \times 10^{-8}$ | 1,564 | Meta-analysis<br>RAPPER-II, N=1,255<br>GenePARE-II, N=161 | -<br>0.98<br>0.98 | 1.37 (1.01 to 1.86)<br>1.27 (0.88 to 1.83)<br>1.61 (0.92 to 2.82) | 0.05 |
| rs1801516 *ATM* | chr11:108,175,462 | A | 0.22 | Overall toxicity[‖,¶] | 1.21 (0.98, 1.49) | NR | 2,697 | Meta-analysis<br>RAPPER-II, N=859<br>GenePARE-II, N=101<br>CCI-BT, N=82 | -<br>NA[#]<br>NA[#]<br>NA[#] | 1.37 (1.05 to 1.78)<br>1.36 (1.03 to 1.80)<br>1.45 (0.63 to 3.34)<br>1.18 (0.21 to 6.55) | 0.02 |
| rs7582141 *TANC1* | chr2:159,899,489 | T | 0.02 to 0.05[††] | Overall toxicity[¶] | 6.17 (2.25, 16.9) | $4.2 \times 10^{-10}$ | 1,742 | Meta-analysis<br>RAPPER-II, N=1,340<br>GenePARE-II, N=220<br>UGhent, N=285<br>CCI-BT (N=114)<br>CCI-EBRT, N=148 | -<br>0.96<br>0.96<br>0.96<br>0.96<br>NA[#] | 0.98 (0.52 to 1.86)<br>0.56 (0.20 to 1.59)<br>0.85 (0.20 to 3.67)<br>2.16 (0.71 to 6.53)<br>NA[**]<br>0.73 (0.08 to 6.38) | 0.95 |

[*] Imputation info score values in CCI-EBRT are from the AffySNP6.0 array; values in all other studies are from the oncoarray. Abbreviations: SNP, single nucleotide polymorphism; OR, odds ratio; CI, confidence interval; MAF, minor allele frequency; OR, odds ratio; NA, not analyzed; NR = not reported.

[†] Base position according to Genome Reference Consortium Human Build 37 (hg19).

[‡] Minor allele frequency is from PRACTICAL Oncoarray samples of European ancestry.

[§] Two-sided $P_{meta}$ was calculated using a Wald test.

[‖] The previously published study included both acute and late toxicity whereas the current study only includes late toxicity.

[¶] Overall toxicity was measured using STAT score [21] based on the worst toxicity grade from 2 years to 5 years after the start of radiotherapy. Analysis is adjusted for pre-radiotherapy STAT score, age, androgen deprivation therapy, prostatectomy, and total BED. Analysis in RADIOGEN used genotype data from the Illumina Oncoarray while the previously published results used genotype data from the Affymetrix Axiom Genome-Wide CEU 1 array [22]. Additional toxicity follow-up data was available in the current analysis that was not available in the earlier analysis.

[#] SNP was directly genotyped.

[**] STAT score was not assessed in CCI-BT because it correlated perfectly with rs7582141 genotype.

[††] The minor allele frequency for rs7582141 and other SNPs in this locus vary across European sub-populations. The frequency of the C allele is 0.024 in RAPPER-I, 0.022 in RAPPER-II, 0.039 in RADIOGEN, 0.042 in GenePARE-I, 0.046 in GenePARE-II, 0.029 in UGhent, and 0.033 in CCI-EBRT.

**Table 4**. Multivariable models including SNPs and clinical risk factors. All models are stratified by study.

| Model | HR (95% CI) | p-value[*] |
|---|---|---|
| Rectal Bleeding[†] | | |
| rs17055178 | 1.84 (1.49 to 2.24) | <0.001 |
| Rectum volume (cm$^3$) receiving 65Gy[‡] | 1.33 (1.08 to 1.63) | 0.007 |
| Rectum volume (percent) receiving 70Gy[§] | 1.44 (1.18 to 1.77) | <0.001 |
| Arthritis | 2.06 (1.12 to 3.48) | 0.02 |
| Inflammatory bowel diverticular disease | 1.80 (1.07 to 2.83) | 0.03 |
| Rectal dose standard deviation[§] | 1.10 (1.03 to 1.18) | 0.008 |
| Intestinal volume (percent) receiving 15Gy[¶] | 1.26 (1.03 to 1.52) | 0.03 |
| Gleason score ≥ 7[#] | 1.25 (1.00 to 1.57) | 0.05 |
| Cardiovascular disease | 1.44 (1.01 to 2.02) | 0.05 |
| Increased Urinary Frequency[**] | | |
| rs17599026 | 1.37 (1.08 to 1.71) | 0.01 |
| Age at treatment > 75[††] | 1.50 (1.16 to 1.92) | 0.002 |
| Diabetes | 1.53 (1.15 to 2.00) | 0.005 |
| Cardiovascular disease | 1.57 (1.04 to 2.31) | 0.04 |
| Prior pelvic surgery | 1.57 (1.06 to 2.24) | 0.02 |
| Presence of hemorrhoids | 1.56 (1.02 to 2.27) | 0.04 |
| Decreased Urinary Stream[‡‡] | | |
| rs10969913 | 2.23 (1.36 to 3.44) | 0.002 |
| rs7720298 | 1.25 (1.05 to 1.48) | 0.01 |
| Presence of hemorrhoids | 2.06 (1.29 to 3.13) | 0.004 |
| Prior TURP | 1.67 (1.13 to 2.39) | 0.01 |
| Bladder volume (cm$^3$) receiving 70Gy[§§] | 1.35 (1.09 to 1.87) | 0.002 |
| Hematuria[‖‖] | | |
| rs11122573 | 1.77 (1.39 to 2.23) | <0.001 |
| rs75991123[¶¶] | 1.61 (1.22 to 2.09) | <0.001 |
| Prior TURP | 2.33 (1.70 to 3.12) | <0.001 |
| Bladder volume (percent) receiving 74Gy[##] | 1.29 (1.09 to 1.51) | 0.003 |
| Receipt of EBRT[***] | 1.92 (1.17 to 3.20) | 0.01 |
| Age at treatment[†††] | 2.80 (1.21 to 5.91) | 0.02 |

Abbreviations: HR, hazard ratio; TURP, transurethral resection of the prostate; EBRT, external beam radiotherapy

[*] Two-sided p-value was calculated using a Wald test.
[†] There were only 6 rectal bleeding events in UGhent and so this cohort was excluded from the model.
[‡] Variable was log2 transformed and includes a spline knot at 3.0cm$^3$, the 25th percentile value. Hazard ratio is per doubling of volume above the 25th percentile value, with reference being values below the 25th percentile.
[§] Variable was log2 transformed and includes a spline knot at 1.7 percent, the 75th percentile value. Hazard ratio is per doubling of percent above the 75th percentile, with reference being values below the 75th percentile.
[‖] This variable is defined as the standard deviation from the mean rectal dose for the standardized rectal volume defined as a solid organ, for each individual patient's dosimetry. It includes a spline knot at 19.7Gy, the median value. Hazard ratio is per unit above the median value, with reference being values below the median.
[¶] Variable was log2 transformed and includes a spline knot at 3.3 percent, the 75th percentile value. Hazard ratio is per doubling of percent above the 75th percentile, with reference being values below the 75th percentile.
[#] Reference group is Gleason < 7.

35

** There were only 8 increased urinary frequency events in UGhent and so this cohort was excluded from the model.

†† Reference group are men ≤ 75 at time of treatment.

‡‡ There were only 12 decreased urinary stream events in CCI-EBRT and so this cohort was excluded from the model.

§§ Variable was log2 transformed. Hazard ratio is per doubling of volume.

‖‖ There were only 6 hematuria events in CCI-EBRT and so this cohort was excluded from the model.

¶¶ The top SNPs in the second region associated with hematuria, rs147121532, has a minor allele frequency < 4% and so the next most strongly associated SNP was used in the multivariable model (minor allele frequency 6%).

## Variable was log2 transformed and includes a spline knot at 1.9 cm$^3$, the median value. Hazard ratio is per doubling of volume above the median value, with reference being values below the median. In UGhent, bladder volume (percent) receiving 75Gy was used instead of bladder volume (percent) receiving 74Gy.

*** Reference group is receipt of brachytherapy alone.

††† Age is treated as a continuous variable if above 75 years. Hazard ratio is per year of age above 75 years, with reference being men with age less than or equal to 75.

**Figure titles and legends**

**Figure 1. Cumulative probability of radiotoxicity.** Each graph shows the cumulative probability of developing grade 2 or worse radiotoxicity for each individual outcome within each study included in the GWAS meta-analysis. These outcomes include rectal bleeding (A), increased urinary frequency (B), decreased urinary stream (C), and hematuria (D). Numbers listed below the x-axis for each graph represent the numbers of patients at risk.
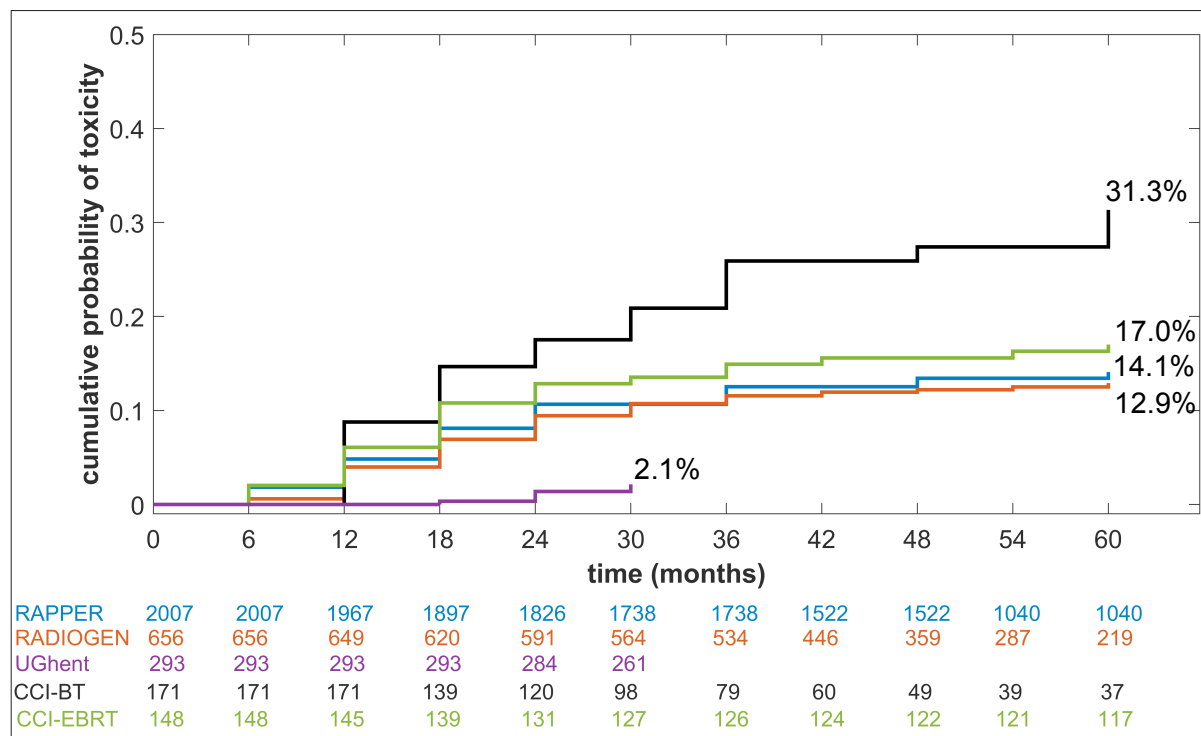
**Figure 2. Manhattan plots.** The graphs show association results for rectal bleeding (A), increased urinary frequency (B), decreased urinary stream (C), and hematuria (D). The red line denotes -log p value = $5\times10^{-8}$. Each point represents a SNP, with numbers on the x-axis denoting chromosome number.

**Figure 3. Regional Manhattan plots.** The graphs show signals defined by fine-mapping of the hematuria risk region chr1:230337180-231337180 (A), rectal bleeding risk region chr5:156903410-157903410 (B) and decreased urinary stream risk region chr9:30366808-31366808 (C).
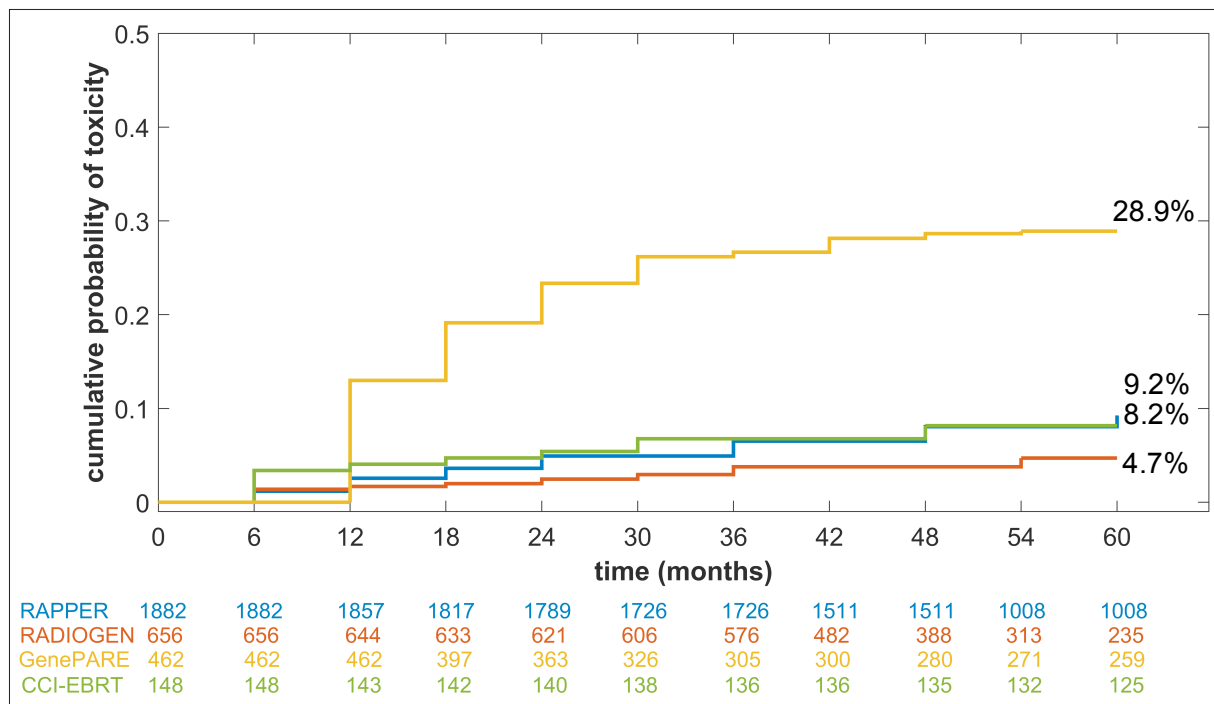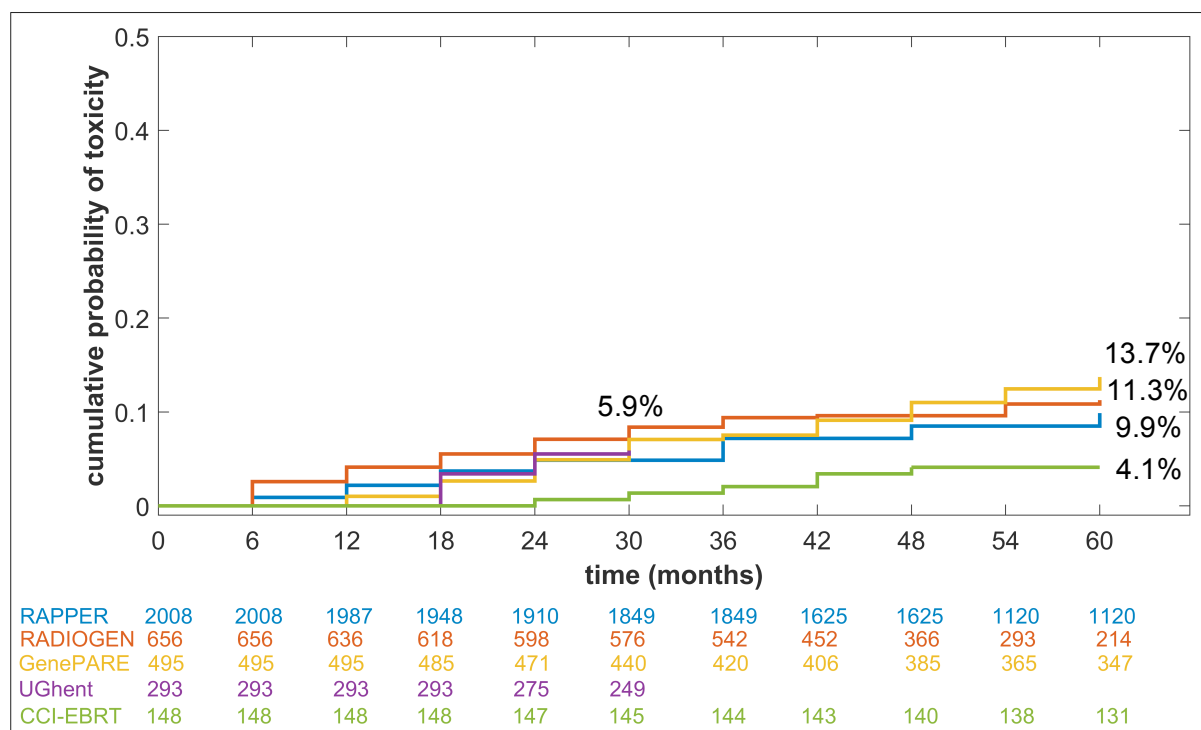
37

Figure 1--FINAL

**Figure 1.**

Figure 2--FINAL

**Figure 2.**

Figure 3--FINAL

**Figure 3**.