Big-RT: Big Data Analysis to Identify Combinatorial Predictors of Radiotherapy Toxicity for Personalised Treatment in Prostate Cancer Patients

Sebastian Pölsterl, Anna Wilkins, Sarah Gulliford, James Campbell, Carmen Rodriguez-Gonzalvez, Sheng Yu, Veronica Garcia-Perez, Claire Griffin, Judith Bliss, John Yarnold, Uwe Oelfke, David Dearnaley, Emma Hall, Bissan Al-Lazikani, Navita Somaiah

Background

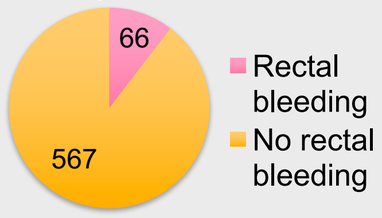
- ~ ⅔ of cancer patients will receive radiotherapy (RT).
- Up to 20% experience **long-term RT-induced toxicity**.
- **Problem:** no stratification by risk of side effects, which limits the ability to personalise treatment.
- **Challenge:** The pathogenesis of RT-induced toxicity is complex and multi-factorial, yet most **prior work focused on isolated data types** and standard statistical techniques, with limited success.

Aim

To identify multi-parametric predictors of *late* RT-induced rectal bleeding by developing novel machine learning techniques that **fully integrate** clinical, dosimetric, and genetic data from the CHHiP prostate RT-fractionation trial (CRUK/06/016)[1,2]**.**

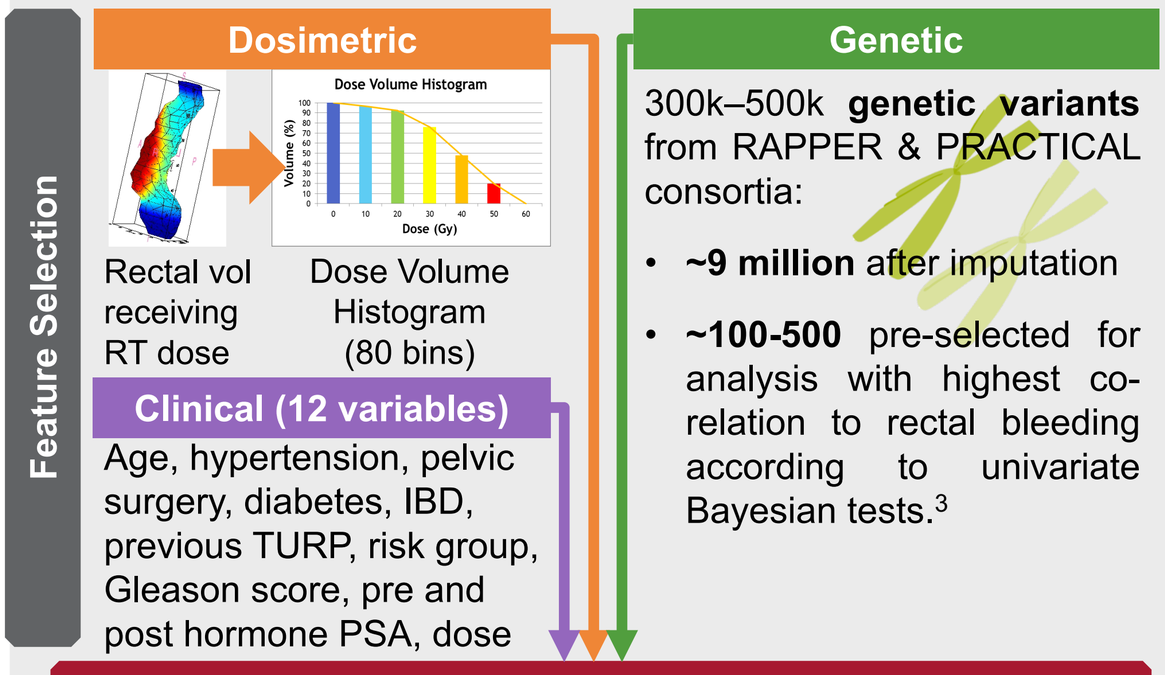
Data

- **3,212** patients recruited → **721** have all data types.
- 88 patients excluded if rectal bleeding information is missing for >3 time points and/or had heamorroids and/or rectal bleeding prior to radiotherapy.
- **633 patients eligible for analysis.**
- Rectal bleeding defined as having grade ≥ 2‡ @t ≥ 12 months.



‡: Moderate rectal bleeding (simple out-patient management)

Methods



Results

Fig 1. *Top*: Median (AUROC) performance scores for models using clinical, genetic and dosimetric factors. Black lines indicate the range of scores for our 10-fold cross validation. *Bottom*: Results using subsets of the data and **Naïve Bayes**, **our best performing model. The strongest predictions are achieved by jointly analysing all data types.**

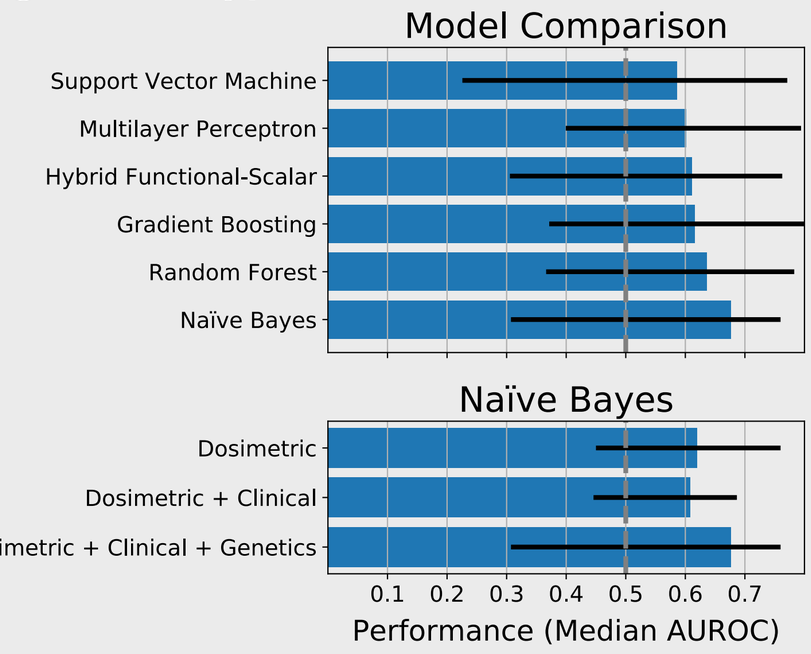
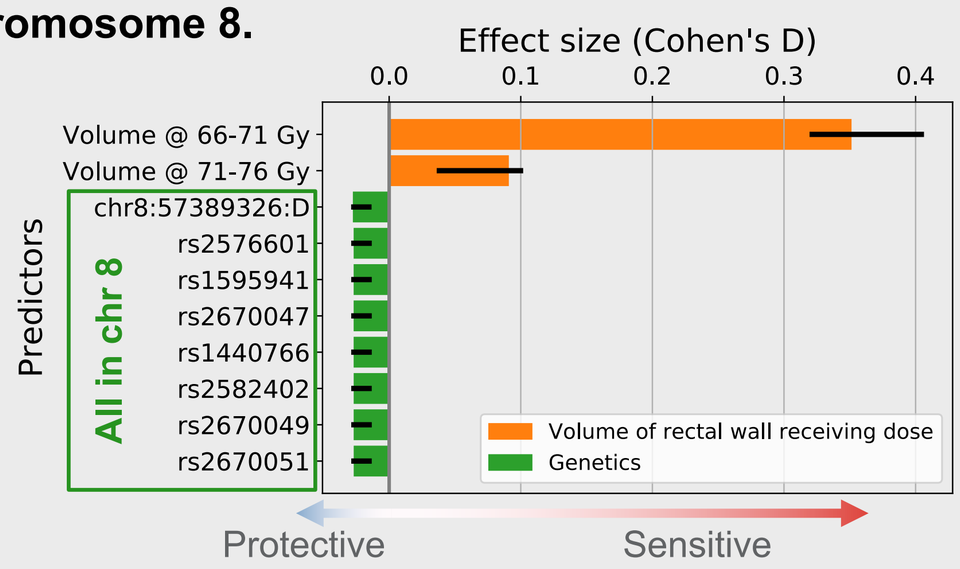


Fig 2. Strength of association to late RT-induced rectal bleeding for our top 10 predictors for the Naïve Bayes model using all data types. Genetic features shown were selected in at least 9 out of 10 folds. **Higher risk of rectal bleeding was associated with an increase in rectal volume receiving 66-76Gy, whilst lower risk was linked to genetic variants in a region of chromosome 8.**



Conclusions

- Identified novel combinatorial markers predictive of RT-induced rectal bleeding.
- Markers must be validated on independent dataset.
- Techniques developed here could be applied to other tumour types treated with radiotherapy.

References

1. Dearnaley et al. Lancet Oncol 2016; 17: 1047–60.
2. Wilkins et al. Lancet Oncol 2015; 16: 1605–16.
3. Marchini et al. Nature Genet 2007; 39: 906–913.
4. Ramsay and Silverman. Functional Data Analysis. Springer. 2005.
5. Shahriari et al. Proc IEEE. 2016 Jan;104(1):148–75.

Contact

The Institute of Cancer Research
The Knowledge Hub
15 Cotswold Road, Sutton, Surrey SM2 5NG, UK
sebastian.poelsterl@icr.ac.uk

Acknowledgments

We acknowledge NHS funding to the NIHR Biomedical Research Centre at The Royal Marsden (RM) and The Institute of Cancer Research (ICR). This work has been supported by the CRUK Centre at the ICR and RM (C309/A18077). the CRUK CTU programme grant (C1491/A15955), The Heather Beckwith Charitable Settlement, The John L Beckwith Charitable Trust, and the ICR Team Science award. We thank Julia Murray, Katherine Fernandez, and Richard Boyle for processing of the dosimetric data, and Catherine West and RAPPER/PRACTICAL consortia for the GWAS data.