The Institute of Cancer Research

PHD STUDENTSHIP PROJECT PROPOSAL

PROJECT DETAILS	
Project Title:	Building prediction statistical models to characterise the dynamics of clinical and pathological features and patient outcome of localised prostate cancer prognosis after external beam radiotherapy
Short Project Title:	Dynamic prediction statistical models to characterise localised prostate cancer prognosis

SUPERVISORY TEAM	
Primary Supervisor(s):	PROF EMMA HALL
Other supervisory team members:	DR NURIA PORTA DR ALISON TREE DR CHRIS PARKER PROF DAVID DEARNALEY

DIVISIONAL AFFILIATION	
Primary Division:	CLINICAL STUDIES
Primary Team:	CLINICAL TRIALS AND STATISTICS UNIT – Prof Emma Hall

PROJECT PROPOSAL

BACKGROUND TO THE PROJECT

Prostate cancer is the most common cancer in men in the UK, with 46,690 new cases in 2014 (CRUK 2017). Initial management of localised disease includes external beam radiation therapy (EBRT) or radical prostatectomy. The majority of patients in the UK receive EBRT. Following radiation, disease is monitored for recurrence using the blood test prostate specific antigen (PSA), which is repeatedly measured over time. Rises in PSA suggest the disease is re-growing, and trigger further investigations. To more accurately guide clinical decision making, monitoring of PSA after EBRT would be aided by dynamic prognostic tools that incorporate the complete post-treatment PSA evolution, amidst other disease-related information.

Prospective long follow-up and high quality data in clinical trials provide a perfect opportunity to explore different patterns of prognosis of disease. Measurements for different outcomes are collected for each patient, and each outcome usually analysed separately. However, joint statistical modelling of different outcomes and available clinical and pathological markers collected alongside follow-up would allow more accurate predictions of the prognosis of a patient. The long follow-up permits a better understanding of how patients transition through different, progressive stages of their disease.

We will develop prognostic models based on data from 2 large practice-changing clinical trials investigating radiotherapy interventions in localised prostate cancer. The MRC RT01 trial (Dearnaley et al., 2014) showed that escalated-dose conformal radiotherapy with short-course neoadjuvant androgen deprivation therapy (ADT) improved biochemical progression-free survival. The CHHiP trial (Dearnaley et al., 2016) showed that hypofractionated radiotherapy using 60 Gy in 20 fractions is non-inferior to conventional fractionation using 74 Gy in 37 fractions in patients receiving short-course ADT and has been

recommended as standard of care for EBRT in this patient population. Access to the trials data will be accessed with the permission of the respective Trial Management Groups.

PROJECT AIMS

The overarching aim of the project is to develop dynamic prediction statistical models for disease prognosis in localised prostate cancer patients following EBRT and neoadjuvant ADT, to obtain personalised predictions enriched by integrating both pre-EBRT information (time-fixed or baseline) and time-dependent information collected during follow-up.

This will be achieved through the following specific aims:

- To describe prostate cancer prognosis dynamics as observed in the MRC RT01 and CHHiP trials by identification of time-dependent events and longitudinal processes of interest.
- Development and validation of dynamic prediction models for the prognosis of prostate cancer incorporating time-dependent information.
- Development of a user-friendly web-based risk calculator implementing the models to obtain individual predictions of prognosis that would help clinical management of future patients.

RESEARCH PROPOSAL

Research plan:

1) Literature review on prostate cancer and statistical models for dynamic prediction:

The candidate will undertake a literature review to understand the diagnosis, management and follow-up of localised prostate cancer patients. In the follow-up of prostate cancer patients after EBRT, a patient can experience localised or distant recurrence; PSA failure, or recommence hormone therapy as palliative therapy following a rise in PSA levels; he can die of prostate cancer or other causes. Other non-disease related events such as a second primary may alter the prognosis of the patient. All these events are detected through careful monitoring and follow-up (collection of PSA and other blood markers at regular time points, imaging and symptoms' assessments).

The candidate will also undertake a contemporary literature review of statistical methods for dynamic prediction appropriate to be used in the context of these trials. These statistical models steam from survival analysis' and longitudinal data' methodologies, and include **multi-state models** (Andersen and Keiding, 2002, Putter et al. 2007), **linear mixed models** (Verbeke and Molenberghs, 2000) and **joint models** (combining longitudinal and time to event components) (Tsiatis and Davidian, 2004, Ibrahim et al., 2010).

2) Description of the dynamics of disease prognosis in localised prostate cancer:

The candidate will apply the appropriate statistical models to data from the MRC RT01 and CHHIP trials, in order to describe and understand the different aspects of disease prognosis, including:

- Modelling the association between PSA failure, disease recurrence (local, nodal or distant), death and other disease-related events such as re-commencement of long-term salvage ADT.
- Modelling of the PSA trajectories after short-course ADT with time, and how these trajectories change following disease-related events.
- Modelling the association between PSA trajectories and disease relapse; identification of class PSA trajectories with different disease prognosis.

3) Development and validation of prognostic models for dynamic prediction:

The above approaches will be considered in order to develop our final prognostic models for dynamic prediction. Combining different dynamic processes, though in theory appealing, may increase complexity of the models and the estimation process, and much information is needed to obtain a desired level of precision to assess their predictive performance. We will consider those approaches that, while increasing our knowledge of the disease prognosis, do not hamper its utility by overcomplicated modelling.

Dynamic predictions for a disease-related event of interest are obtained by computing the probability or risk of the event occurring by time *s+t*, given the history of the patient up to time *s.* Such history includes any information on the patient available at time s: baseline covariates, PSA values up to time s, other biomarkers available, the occurrence or absence of related events such as patient on ADT treatment or not. Such predictions can be obtained from the above models (Sene et al., 2016). For instance, to predict the probability of being free of biochemical or clinical recurrence in 5 years, for a patient with intermediate risk, with Gleason 7, T2a stage and pre-ADT PSA of 15 ng/mL (baseline features), AND knowing that, 2 years after EBRT, he has not re-commenced ADT, he is free of local or distant recurrence, and his PSA values at 6, 12 and 15 months are 2.1, 2.2 and 3.2 ng/mL (time-dependent information).

Predictive performance measures to assess calibration and discrimination will be used to compare and select the best models. Validation of dynamic predictions can be challenging, as are usually conditioned on scenarios of initiation (i.e. what is the history of the patient at time s) and often assume these conditions remain fixed from *s* to *s+t* - validation of the predictive performance under several scenarios is required.

The initial strategy will be to use data from the CHHiP and MRC RT01 trials to develop the models and ensure that sufficient numbers of disease events are available; internal validation will be performed using bootstrapping or cross-validation techniques, and suitable independent cohorts for external validation will be sought.

In addition, the model may also be applied to data from the PACE-B trial (CRUKE/12/025) to see if it is generalizable to patients who did not receive short-course of neoadjuvant RT, and other hypofractionation regimes. PACE-B is a trial coordinated by ICR-CTSU currently open to recruitment, which will randomise 858 prostate cancer patients to either EBRT (clinician can choose hypofractionated regime) or stereotactic body radiotherapy (SBRT) given in 5 fractions.

4) Development of an online web-based risk calculator to obtain individual predictions of disease prognosis:

This tool will aim to disseminate our prognostic model for clinical use, and aid the clinicians with decision making in the management of future patients with localised prostate cancer. The candidate will be involved in the development of this web-based risk calculator or nomogram, using the statistical software R (https://cran.r-project.org/), and potentially its Shiny package (https://shiny.rstudio.com/), which builds interactive web applications straight from R. Consideration to other software (e.g. STATA) may be given to overcome computational issues. Examples of prostate cancer nomograms have been developed in the past (for instance to predict outcomes after radical prostatectomy https://www.mskcc.org/nomograms/prostate) but often these nomograms are based on baseline information and do not allow for time-dependent information, such as post-RT PSA values.

<u>Projected outputs:</u>

Peer-reviewed manuscripts are anticipated in both clinical and methodological journals. A risk calculator to be available for the clinical community as an online web-based resource.

LITERATURE REFERENCES

ANDERSEN, P. K. & KEIDING, N. 2002. Multi-state models for event history analysis. Stat Methods Med Res, 11, 91-115.

CANCER RESEARCH UK. 2017. Prostate cancer statistics - key facts [Online]. http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/prostate-cancer [Accessed August 2017].

DEARNALEY, D., SYNDIKUS, I., MOSSOP, H., KHOO, V., BIRTLE, A., BLOOMFIELD, D., GRAHAM, J., KIRKBRIDE, P., LOGUE, J., MALIK, Z., MONEY-KYRLE, J., O'SULLIVAN, J. M., PANADES, M., PARKER, C., PATTERSON, H., SCRASE, C., STAFFURTH, J., STOCKDALE, A., TREMLETT, J., BIDMEAD, M., MAYLES, H., NAISMITH, O., SOUTH, C., GAO, A., CRUICKSHANK, C., HASSAN, S., PUGH, J., GRIFFIN, C., HALL, E. & INVESTIGATORS, C. H. 2016. 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, 1047-60.

DEARNALEY, D. P., JOVIC, G., SYNDIKUS, I., KHOO, V., COWAN, R. A., GRAHAM, J. D., AIRD, E. G., BOTTOMLEY, D., HUDDART, R. A., JOSE, C. C., MATTHEWS, J. H., MILLAR, J. L., MURPHY, C., RUSSELL, J. M., SCRASE, C. D., PARMAR, M. K. & SYDES, M. R. 2014. Escalated-dose versus control-dose conformal radiotherapy for prostate cancer: long-term results from the MRC RT01 randomised controlled trial. Lancet Oncol, 15, 464-73.

IBRAHIM, J. G., CHU, H. & CHEN, L. M. 2010. Basic concepts and methods for joint models of longitudinal and survival data. J Clin Oncol, 28, 2796-801.

PUTTER, H., FIOCCO, M. & GESKUS, R. B. 2007. Tutorial in biostatistics: competing risks and multi-state models. Stat Med, 26, 2389-430.

SENE, M., TAYLOR, J. M., DIGNAM, J. J., JACQMIN-GADDA, H. & PROUST-LIMA, C. 2016. Individualized dynamic prediction of prostate cancer recurrence with and without the initiation of a second treatment: Development and validation. Stat Methods Med Res, 25, 2972-2991.

TSIATIS, A. A. & DAVIDIAN, M. 2004. Joint modeling of longitudinal and time-to-event data; an overview. Statistica Sinica, 14, 809-834.

VERBEKE, G. & MOLENBERGHS, G. 2000. Linear Mixed Models for Longitudinal Data, Springer.

CANDIDATE PROFILE

Note: the ICR's standard minimum entry requirement is a relevant undergraduate Honours degree (First or 2:1)

Pre-requisite qualifications of applicants: e.g. BSc or equivalent in specific subject area(s)	MSc or BSc (2:1 or better) in Mathematical Sciences including a statistical component or Statistics

Intended learning outcomes:

Please provide a bullet point list of the knowledge and skills you expect the student to have attained on completion of the project.

Ability to understand and critically leverage complex statistical models that best fit our clinical problem

Ability to write abstracts and papers on this and related subjects

Ability to implement complex statistical analysis in R/Stata