and showed to be associated with toxicity. It is important to remember that such features, to some extent, might be confounded by more simple factors (e.g. tumor volume or volume of irradiated region). Nevertheless, image based features appears in a number of studies to add independent toxicity information; but it is likely that no single image-based feature (or no single feature at all) will be able to make a perfect patient specific toxicity prediction for the entire population. In many studies the correlation between a specific image-based feature and observed toxicity is relative weak. However, if predictive toxicity models simply are able to identify a subset of patients who are likely to have modest toxicity that would be very beneficial, since this group of patients could then be offered a more aggressive treatment, which hopeful would result in improved local control. Predictive toxicity models should thus not only be evaluated on their overall prediction performance for the entire population, but also on their ability to identify a significant subgroup of patients who are candidates for intensified treatment.
The current lecture will present examples of image-based features and point to their potential clinical impact; but will also focus on the potential use of patient specific toxicity models to select subgroups of patients as described above. Moreover comments on image quality will be made, since high images quality is the foundation for imaged-based features used in predictive models for toxicity.

SP-0310
Growing importance of data-mining methods to select dosimetric/clinical variables in predictive models of toxicity
T. Rancati[1]
[1]*Fondazione IRCCS Istituto Nazionale dei Tumori, Prostate Cancer Program, Milan, Italy*

In the field of toxicity modeling it is common practice to build statistical models starting from analysis of clinical data which are prospectively collected in the frame of observational trials. Modern prospective observational studies devoted to modelling of radioinduced toxicity are often accumulating a large amount of dosimetric and patient-related information, this requires particular attention when normal tissue complication probability modelling is approached. A core issues is related to selection of features, which then influences overfitting, discrimination, personalization and generalizability.
These risks are particularly high in clinical research datasets, which are often characterized by low cardinality - i.e. the number of cases is overall low - and are often strongly imbalanced in the endpoint categories – i.e. the number of positive cases (e.g. toxicity events or loss of disease control) is small, or even very small, with respect to the negative ones. This is obviously positive for patients, it is however a disadvantage for model building.
In this context a possible methods using in-silico experiment approach for toxicity modelling will be discussed together with some applications.
This method aimed at identifying the best predictors of a binary endpoint, with the purpose of detecting the leading robust variables and minimizing the noise due to the particular dataset, thus trying to avoid both under- and over-fitting. It followed, with adjustments, a procedure firstly introduced by El Naqa [IJROBP2006]: the treatment response curve was approximated by the logistic function, while the bootstrap resamplings were performed to explore the recurrence of the selected variables in order to check their stability. A further bootstrap resampling was introduced for the evaluation of the odds ratios of the selected variables.
The in-silico experiment was implemented using the KNIME software (KNIME GmbH, Germany) and consisted in the following processing steps:
1) 1000 bootstrap samplings of the original dataset are created, as suggested by El Naqa [IJROBP2006];
2) backward feature selection based on minimization of residuals is performed on each bootstrap sample;
3) the rate of occurrences and the placement of each variable (selected by the backward feature selection) in the 1000 bootstrapped datasets are used to classify the most robust predictors. A synthetic index, called normalized area, is defined for ranking each predictor: it corresponds to the area under the histogram representing the number of occurrences of each variable (x-axis) at a given importance level in each re-sampled dataset;
4) a basket analysis of the 1000 sets of predictors is used to identify the predictors that appears together with higher probability;
5) the best set of predictors is chosen, with its maximum size determined by the rule of thumb “one tenth of the number of toxicity events”;
6) the distribution of odds ratios are determined through 1000 bootstrap re-samplings of the original dataset including the set of predictors selected in the previous step;
7) a logistic model with the best set of predictors and the median odds ratios, calculated from the distributions obtained in the previous step, is defined.
In this approach, logistic regression is enhanced with upstream and downstream data processing to find stable predictors.
The method was tested with satisfactory results on different datasets aimed at modelling radio-induced toxicity after high-dose prostate cancer radiotherapy.

Symposium: Automated treatment plan generation in the clinical routine

SP-0311
Automated treatment plan generation - the Zurich experience
J. Krayenbuehl[1], M. Zamburlini[1], I. Norton[2], S. Graydon[1], G. Studer[1], S. Kloeck[1], M. Guckenberger[1]
[1]*University Hospital Zürich, Department of Radiation Oncology, Zurich, Switzerland*
[2]*Philips, Philips Radiation Oncology Systems, Fitchburg, USA*

Intensity modulated radiotherapy and volumetric modulated radiotherapy (VMAT) involves multiple manual steps, which might influence the plan quality and consistency, for example planning objectives and constraints need to be manually adapted to the patients individual anatomy, tumor location, size and shape [1]. Additional help structures are frequently defined on an individual basis to further optimize the treatment plan, resulting in an iterative process. This manual method of optimization is time consuming and the plan quality is strongly dependent on planner experience. This is especially true for complex cases such as head and neck (HN) carcinoma and stereotactic treatment.
In order to improve the overall plan quality and consistency, and to decrease the time required for planning, automated planning algorithms have been developed [2,3]. In this pilot study, we compared two commercially available automatic planning systems for HN cancer patients. A VMAT model was created with a knowledge based treatment system, Auto-Planning V9.10 (Pinnacle, Philips Radiation Oncology Systems, Fitchburg, WI) [4] and for a model based optimization system, RapidPlan V13.6 (Eclipse, Varian Medical System, Palo Alto, CA) [2]. These two models were used to optimize ten HN plans. Since the aim was to achieve plans of comparable quality to the manually optimized plans in a shorter time, only a single cycle of plan optimization was done for both automated treatment planning systems (TPS). Auto-Planning was additionally used to evaluate the treatment of lung and brain metastases stereotactic treatments.
The results from the planning comparison for HN cancer patients showed a better target coverage with AutoPlanning in comparison to Rapidplan and manually optimized plans ($p < 0.05$). RapidPlan achieved better dose conformity in comparison to AutoPlanning ($p < 0.05$). No significant differences were observed for the OARs, except for the swallowing muscles where RapidPlan and the manually optimized plans were better than AutoPlanning and for the mandibular bones were AutoPlanning performed better than the two other systems. The working time needed to generate