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

clinical accepted plans for both automated TPS was drastically reduced to less than ten minutes.

For the two stereotactic sites evaluated, target coverage and OARs doses differences were not clinically relevant between Auto-Planning and manually optimized plans.

The encouraging results of automatic planning shows that highly consistent treatment plans for complex cases can be achieved with an automated planning process.

SP-0312
Automated treatment plan generation - the Milan experience
A. Fogliata[1]
[1]Humanitas Research Hospital, Department of Radiation Oncology, Rozzano-Milan, Italy

A knowledge based planning process, named RapidPlan, has been recently implemented into the Varian Eclipse treatment planning system. The goal of the engine is to generate patient tailored and personalized objectives to input in the optimization process for IMRT or VMAT inverse planning. Data from previously generated high quality plans are used to estimate DVH ranges where the specific DVH of a structure will most likely land according to the prior plans knowledge. Estimate-based optimisation objectives are hence generated. A complete pre-clinical preparation have been established before the clinical implementation of RapidPlan and the configured specific models. The anatomical sites and pathologies chosen for the first models generation in Milan were Head and Neck, and Breast. For the first site the choice was driven by the complexity of the planning phase due to the anatomy and critical structures; the breast was chosen since, beside of its planning complexity, almost one third of our patient population presents breast cancer. For each of the two chosen sites the process of the model generation included different phases. Initially a set of about 100 patients per site, having quite spread anatomical characteristics (as, for example, the breast size) while excluding extreme anatomies, was selected. The selected plans were all clinical plans of high quality, for VMAT (RapidArc) delivery. Those plans were used to train the model for the extraction of the parameters, based on prinicipal component analysis methods and regression models, needed to estimate the DVH for any new patient. The training results were analysed to evaluate possible outliers and their eventual exclusion from the model. Finally the validation process was followed on another group of patients to assess the model reliability and usability. From this last phase improvements in the plan quality when using RapidPlan was assessed. Once the two models were evaluated, a number of head and neck and breast cases were selected for the pre-clinical trial. The planners used to plan without RapidPlan were asked to produce plans using the knowledge based planning models. Two kind of evaluations were felt interesting: on one side the plan quality, for which the same cases were asked to be planned without RapidPlan by the same planner, and on the other side the time required to obtain such plans. The results were very promising, both on the plan quality, and especially on planning time. We are ready to move to the clinical daily use of the automated treatment plan generation.

SP-0313
Fully automated treatment plan generation using Erasmus-iCycle - the Rotterdam experience
M.L.P. Dirkx[1], A.W. Sharfo[1], P.W.J. Voet[2], G. Della Gala[1], L. Rossi[1], D. Fransen[1], J.J. Penninkhof[1], M.S. Hoogeman[1], S.F. Petit[1], A.M. Mendez-Romero[1], J.W. Mens[1], L. Incrocci[1], N. Hoekstra[1], M. Van de Pol[1], S. Aluwini[1], S. Breedveld[1], B.J.M. Heijmen[1]
[1]Erasmus MC Cancer Institute, Radiation Oncology, Rotterdam, The Netherlands
[2]Elekta AB, Physics Research, Uppsala, Sweden

Aim: Treatment plan generation in radiotherapy is commonly a trial-and-error procedure in which a dosimetrist tries to steer the treatment planning system (TPS) towards an acceptable patient dose distribution. For a single patient, this process may take up to several days of workload. The quality of the final treatment plan is dependent on the skills and experience of the dosimetrist, and on allotted time. In addition, for the treating physician it is extremely difficult to assess whether the generated plan is indeed optimal considering the unique anatomy of the individual patient. At Erasmus MC, systems for fully automated plan generation have been developed to obtain plans of consistent high quality, with a minimum of workload. This presentation will focus on their clinical implementation and applications.

Materials and methods: An IMRT or VMAT plan is generated fully automatically (i.e., without human interface) by the clinical TPS (Monaco, Elekta AB), based on a *patient-specific* template. The patient-specific template is automatically extracted from a plan generated with Erasmus-iCycle, our in-house developed pre-optimizer for lexicographic multi-criterial plan generation (Med Phys. 2012; 39: 951-963). For individual patients of a treatment site (e.g., prostate), automatic plan generation in Erasmus-iCycle is based on a *fixed* 'wishlist' with hard constraints and treatment objectives with assigned priorities. The higher the priority of an objective, the higher the chance that the planning aim will be achieved, or even superseded. All plans generated with Erasmus-iCycle are Pareto optimal. In case of IMRT, the system can be used for integrated beam profile optimization and (non-coplanar) beam angle selection. *Site-specific wishlists* are a priori generated in an iterative procedure with updates of the wishlist in every iteration step, based on physicians' feedback on the quality of plans generated with the current wishlist version. Also for patients treated at a Cyberknife, either with the variable aperture collimator (Iris) or MLC, the clinical TPS (Multiplan, Accuray Inc.) can be used to automatically generate a deliverable plan, based on a pre-optimization with Erasmus-iCycle.

Results: Currently, automatic treatment planning is clinically used for more than 30% of patients that are treated in our department with curative intent. It is routinely applied for prostate, head and neck, lung and cervical cancer patients treated at a linac. In a prospective clinical study for head and neck cancer patients, treating radiation oncologists selected the Erasmus-iCycle/Monaco plan in 97% of cases rather than the plan generated with Monaco by trial-and-error (IJROBP 2013; 85: 866-72). For a group of 41 lung cancer patients, clinically acceptable VMAT plans could be generated fully automatically in 85% of cases; in all those cases plan quality was superior compared to manually generated Monaco plans, due to a better PTV coverage, dose conformality, and/or sparing of lungs, heart and oesophagus. For plans that were initially not clinically acceptable, it took a dosimetrist little hands-on time (<10 minutes) to modify them to a clinically acceptable plan. In 44 dual-arc VMAT Erasmus-iCycle/Monaco plans for cervical cancer treatment small bowel V45Gy was reduced by on average 20% (p<0.001) when compared to the plans that were manually generated by an expert Monaco user, spending 3 hours on average. Differences in bladder, rectal and sigmoid doses were insignificant. For 30 prostate cancer patients, differences between Erasmus-iCycle/Monaco VMAT plans and VMAT plans manually generated by an expert planner with up to 4 hours planning hands-on time, were statistically insignificant (IJROBP 2014; 88(5): 1175-9). Attempts to use acceptable, automatically generated plans as a starting point for manual generation of further improved plans have been unsuccessful. For prostate SBRT, clinically deliverable Cyberknife plans that were automatically generated with Erasmus-iCycle/Multiplan showed a better rectum sparing and a reduced low-medium dose bath compared to automatically generated VMAT plans with the same CTV-PTV margin.

Conclusion: In our department, automatic plan generation based on Erasmus-iCycle is currently widely used, showing a consistent high plan quality and a vast reduction in planning workload. Extension to new target sites (breast, liver, lymphoma, spine, vestibular schwannoma) is being investigated. In addition, the use of automated planning for intensity modulated proton therapy is being explored, making objective plan comparison with other modalities possible.