AbstractID: 1424 Title: Normal tissue complication probability modeling techniques using bootstrap replicates of the variable selection process

The probability of an undesirable radiotherapy complication is typically a function of both treatment-related (dose-volume) and clinical factors (such as age, gender, diabetes, etc.). We describe techniques for building NTCP models which include multiple dosimetric and clinical factors, and methods to test the robustness of the metrics (variables) selected for inclusion. Multi-metric logistic regression techniques are used as the modeling basis. Bootstrap replications of Akaike information criteria calculations are used to determine an optimal number of variables without significant over-fitting. Of concern, however, is the robustness of variable selection against inevitable noise in the (limited) data. Put simply, are any of the 'significant' variables produced by the regression process present 'by accident?' To address this ever-present issue, we conduct repeated variable selection and regression using bootstrap data replicates. The results are processed in graphical display, which allows for a greater understanding of correlations between potential model input metrics. Another useful, yet simple, graphical technique is to display the correlation matrix of metrics for the dataset under consideration. This makes clear, for example, the futility of separating out the effects of variables which may be very highly correlated within the dataset under consideration (e.g., V20 and mean lung dose). Analyses from esophagitis, pneumonitis, and xerostomia datasets are used as examples. Multi-metric models, and the bootstrap robustness tests suggested here, can be important tools for improving complication probability models by including dosimetric and clinical factors.

Partially supported by NCI grant CA85181