Conclusion
A comprehensive, radiobiologically consistent Poisson-based TCP model of the response to post-prostatectomy RT was validated for the first time on a completely independent data set. A more extensive validation on a larger population is actually in progress to further corroborate its generalizability.

PO-0853 A method for automatic selection of parameters in NTCP modelling

D. Christophides[1], A.L. Appelt[2], J. Lilley[3], D. Sebag-Montefiore[2]
[1]*Leeds CRUK Centre and Leeds Institute of Cancer and Pathology, University of Leeds, Leeds, United Kingdom*
[2]*Leeds Institute of Cancer and Pathology - University of Leeds and Leeds Cancer Centre, St James's University Hospital, Leeds, United Kingdom*
[3]*Leeds Cancer Centre, St James's University Hospital, Leeds, United Kingdom*

Purpose or Objective
The use of multivariate models in predicting NTCP has the potential of improving predictive accuracy compared to univariate models[1]. However the large numbers of clinical parameters and dose metrics involved can make the selection of the optimal multivariate model inconsistent and time consuming.
In this study a genetic algorithm based method is utilised to automatically generate ordinal logistic regression models; subsequently the quality of the parameter selection process is evaluated by comparison with published results on the same patient cohort[2].

Material and Methods
A general method for selecting optimal models for outcome prediction in radiotherapy was developed (Fig.1). The method was tested on data from 345 rectal cancer patients, used in a previously published study[2], to generate ordinal logistic regression models for the prediction of acute urinary toxicity during chemoradiotherapy. Principal component analysis (PCA) was used to derive principal components (PCs) that summarise the variance in the DVH data. Overall 25 clinical parameters were considered in the analysis including demographics, treatment regime, plan parameters and stage of disease; as well as 8 PCs that explained >95% of the variance in the DVHs.
Urinary toxicity was categorised as grade 0, 1 and 2≥ cystitis, according to the CTCAE v3.0. The method (Fig.1) for optimising the models was implemented in Python and the entire procedure was repeated 100 times, using bootstrap sampling from the whole data set, to evaluate the stability of the parameter selection.
Confidence intervals for the Akaike information criterion (AIC) of the final models selected were estimated using 1000-sample bootstrap.

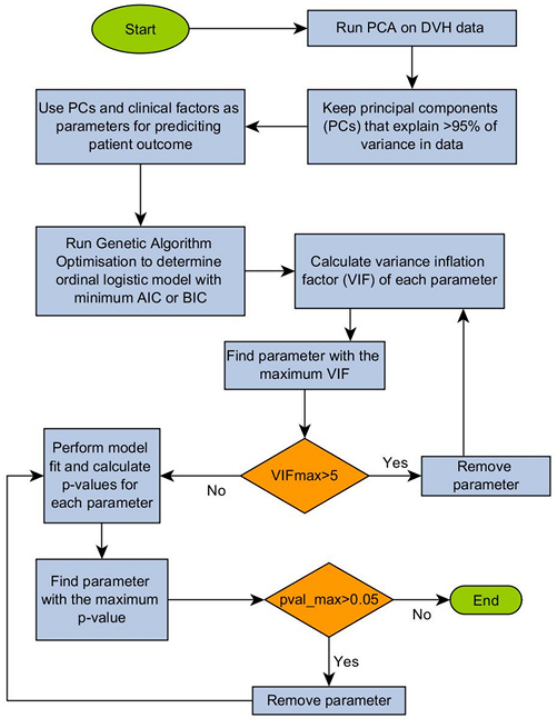

Fig. 1. Flowchart of the method for the automatic generation of ordinal logistic regression models. The variance inflation factor was used to remove collinear parameters following genetic algorithm optimisation; in addition the regression p-values were used to remove non-significant parameters. Two models were produced, one with the genetic algorithm minimising the AIC and another minimising the Bayesian information criterion (BIC).

Results
The method (Fig.1) used to minimise AIC identified PC1, brachytherapy dose level and gender as the optimal model variables. This agreed well with the model identified by Appelt et al[2] that used the $V_{35.4Gy}$, brachytherapy dose and gender; considering that PC1 was found to have a high correlation with the $V_{35.4Gy}$ (R^2=0.96, p<0.001). The model determined by minimising the BIC, identified PC1 and brachytherapy treatment status as important predictive variables. The bootstrap analysis identified PC1 and gender as the most stable parameters.
The 95% bootstrap confidence intervals of the AIC for all three models overlapped significantly; with (625.3, 681.5) for the AIC-minimised model, (627.0, 686.2) for BIC-minimised and (624.8, 680.6) for the published model[2].
The similarity between the models was further demonstrated by plotting the observed and predicted risk with increasing levels of predicted risk (Fig.2).

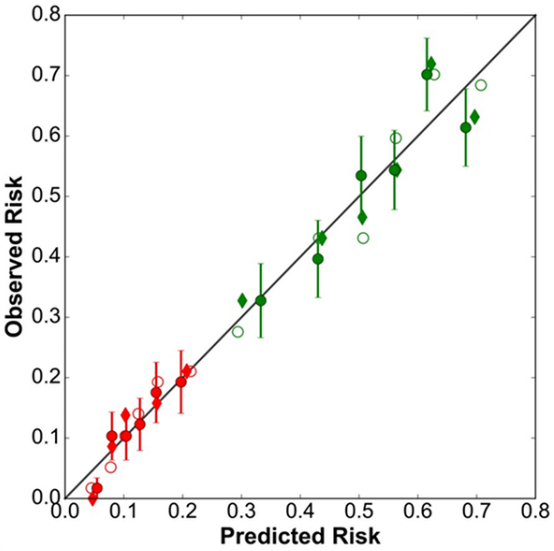

Fig. 2. Scatter plot of predicted vs. observed risk with patients grouped in 6 equally sized bins. Separate predicted risks were calculated, for patients having ≥1 grade cystitis (green) and ≥2 grade cystitis (red). No significant differences can be seen between the published model[2] (empty circles), the model derived by using the genetic algorithm to minimise AIC (solid diamonds) and alternatively minimise BIC (solid circles). Error bars indicate 68% confidence intervals.

Conclusion

The method proposed can automatically generate ordinal logistic regression models that can have equivalent predictive accuracy as models created manually. Furthermore the method can be used to save time in data analysis, tackle problems with a large number of parameters and standardise variable selection in NTCP modelling.

[1] Lind et al (2002) IJROBP 54 340-347
[2] Appelt et al (2014) Acta Oncol. 54 179-186

PO-0854 Is radiation-induced trismus a time dependent masticatory structure story?

M. Thor[1], C. Olsson[2], J. Oh[1], N. Pauli[3], N. Pettersson[4], C. Finizia[3], J. Deasy[1]
[1]*Memorial Sloan Kettering Cancer Center, Department of Medical Physics, NYC, USA*
[2]*Institute of Clinical Sciences- the Sahlgrenska Academy at the University of Gothenburg, Department of Radiation Physics, Gothenburg, Sweden*
[3]*Institute of Clinical Sciences- the Sahlgrenska Academy at the University of Gothenburg, Department of Otorhinolaryngology- Head and Neck Surgery, Gothenburg, Sweden*
[4]*University of California San Diego, Department of Radiation Medicine and Applied sciences, La Jolla, USA*

Purpose or Objective

To investigate temporal radiation-induced etiologies for trismus using dose to five masticatory structures within a thorough internal generalizability approach.

Material and Methods

This study included 93 patients previously treated with primary radiotherapy (RT) for head and neck cancer in 2007-2012 to 64.6-68Gy@1.7-2.0 Gy/fraction. All patients had complete dose data, and trismus assessments (maximum interincisial mouth-opening distance, MIO) at baseline, and at 3, 6, and 12 months post-RT. At each follow-up, the mean dose to each of five masticatory structures (bilateral, contralateral and ipsilateral representations) and ten other patient characteristics was included in a univariate linear regression analysis (UVA) within a 200 times iterated 5-fold cross-validation approach. One additional analysis was performed with the lowest MIO over the three follow-up times as response variable (referred to as "3-12 months"; observed at the 3/6 months follow-ups in 60% of the cases). Candidate predictors from UVA, *i.e.* with a median two-sided p-value≤0.05 over all iterations, qualified for multivariate linear regression analysis (MVA) applying the same cross-validation approach. Predictability was assessed using coefficient of determination (r^2), and Spearman's rank correlation coefficient (Rs); both given as the median over all iterations.

Results

Of 5-12 variables that presented with p≤0.20 on UVA (**Table**), trismus status pre-RT was an independent predictor for post-RT trismus (p=0.01-0.02 for all response variables) as was the mean dose to the ipsilateral masseter (p=0.05 at 3, 6, and 3-12 months). The combination of these two candidate predictors generated MVA models with increased predictability compared to the corresponding UVA models (r^2=0.35-0.40 vs. 0.20-0.32; Rs=0.59-0.63 vs. 0.44-0.57), and consequently steeper response curves with 11-13 mm and 15-16 mm MIO difference between the least and the most risky quintile for the UVA and MVA models, respectively (**Figure**). A tendency of trismus recovery was noted for longer follow-up with a lower pre-RT normalized MIO difference at 12 months compared to that of the two earlier assessments; median (range): 0.14 (-0.67, 0.62) vs. 0.17 (-1.07, 0.66) at 3 months, and 0.16 (-1.33, 0.64) at 6 months.

Table. Predictability for all UVA (p≤0.20; upper), and MVA models for each investigated response variable.

UVA	3 months			6 months			12 months			3-12 months		
Candidate predictor	r^2	*Rs*	*p*	r^2	*Rs*	*p*	r^2	*Rs*	*p*	r^2	*Rs*	*p*
MIO pre-RT	0.32	0.57	0.01*	0.29	0.54	0.02*	0.27	0.52	0.02*	0.27	0.52	0.02*
D_{mean} Mass Ipsi	0.21	0.46	0.05*	0.22	0.47	0.05*	0.20	0.44	0.06	0.21	0.46	0.05*
D_{mean} Mass Bilat	0.20	0.45	0.06	0.20	0.44	0.06	0.19	0.44	0.06	0.18	0.43	0.07
Smoking (0-3)	0.14	0.38	0.12									
D_{mean} PM Ipsi	0.12	0.35	0.15	0.13	0.36	0.13	0.10	0.32	0.19	0.12	0.34	0.16
D_{mean} PM Bilat	0.11	0.33	0.18	0.11	0.34	0.16	0.15	0.38	0.11	0.10	0.32	0.18
D_{mean} TMJ Ipsi				0.10	0.32	0.19	0.12	0.35	0.14			
D_{mean} PL Ipsi							0.13	0.36	0.14			
D_{mean} PL Bilat							0.12	0.35	0.15			
D_{mean} TMJ Bilat							0.11	0.34	0.17			
D_{mean} TMJ Contra							0.11	0.32	0.18			
D_{mean} Mass Contra							0.10	0.32	0.19			
D_{mean} PL Contra							0.10	0.32	0.19			
MVA	**3 months**			**6 months**			**12 months**			**3-12 months**		
Candidate predictor	r^2	*Rs*	*p*	r^2	*Rs*	*p*	r^2	*Rs*	*p*	r^2	*Rs*	*p*
MIO pre-RT + D_{mean} Mass Ipsi	0.39	0.62	0.005*	0.40	0.63	0.004				0.35	0.59	0.008

**Denotes significance, and qualification for MVA if >1 predictor/follow-up time. Abbreviations: Bilat=Bilateral; Contra=Contralateral; D_{mean}=Mean dose; Ipsi=Ipsilateral; Mass=Masseter; PL=Lateral Pterygoid; PM=Medial Pterygoid; TMJ=Temporomandibular joint.*

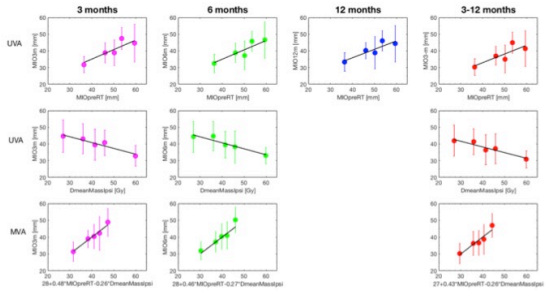

Figure. The UVA dose-response curves for MIO pre-RT (upper), and Dmean Mass Ipsi (middle), as well as the MVA dose-response curves for them combined (lower) for each investigated response variable. *Note: Observed data is given in colored quintiles as the mean (error bars: SD), and predicted as black lines. Abbreviations as in Table.*

Conclusion

A temporally robust dose-response relationship for radiation-induced trismus, quantified as a millimeter mouth-opening decrease, could be observed within the first year after completed radiotherapy. Our results suggest that the dose-response for trismus within this period relies on the mean dose to the ipsilateral masseter, as well as the underlying pre-treatment mouth-opening ability. Up to ten additional variables presented with p-values in the interval p=0.06-0.19 and may prove to be of importance if investigated in larger/pooled cohorts with diversified treatment approaches where potential effects can be thoroughly investigated.