**Head and Neck Quality of Life Study**

**Collected Data:** 225 patients attending head and neck follow-up clinics at the Tom Baker Cancer Centre from June-October 2019 each completed three head and neck patient-reported questionnaires. These questionnaires included: the MD Anderson Symptom Inventory – Head & Neck (MDASI-HN), the MD Anderson Dysphagia Inventory (MDADI), and the Xerostomia Questionnaire (XQ). Data collection was cross-sectional: patients had completed treatment approximately 3 months to 6 years prior to questionnaire completion. Patients provided written consent according to our ethics approval.

In addition to the survey responses for the entire cohort of 225 patients, we have developed a dataset for 156 patients receiving radical (chemo)radiotherapy (70 Gy in 33-35 fxs, 66 Gy in 30 fxs). For these patients, we have collected information from their medical chart, planning CT, treatment plan, on-unit CBCT images, as well as estimated delivered doses. This data is summarized in the table below.

|  |  |  |  |
| --- | --- | --- | --- |
| Patient and Tumour Data from Electronic Medical Record (EMR) | Planning CT Data (pCT) | Treatment Plan Data (RTx) | Patient Monitoring and CBCT-Based Measurements (Obs) |
| Age  Gender  Cancer Site  TNM Stage  Chemotherapy agent  ECOG  Charlson comorbidity index  HPV status  Smoking history  Drinking history  Initial BMI  Disease laterality  Bolus | Structure volumes at planning:   * High-dose CTV * Low-dose CTV * Brainstem * Spinal cord * Pharyngeal constrictor * Ips./cont. parotid gland * Ips./cont. submandibular gland | Planned dose parameter values:   * High-dose CTV D95%, D2% * Low-dose CTV D95%, D20% * Brainstem D0.03cc * Spinal cord D0.03cc * Pharyngeal constrictor Dmean * Ips./cont. parotid gland Dmean * Ips./cont. submandibular gland Dmean | Face diameter  Neck diameter  Neck/shoulder contour  Head rotation  Chin tilt  Shoulder position  BMI  PEG or NG tube placement |

Table: Collected Data for the 156 patients receiving 66-70 Gy in 30-35 fractions.

Ips. = ipsilateral, Cont. = contralateral, = change relative to value at planning

**Potential Study Scope:** Our plan is to produce two studies from this data. The first is to directly report the results of the QoL data collection. Chart reviews may supplement the existing dataset of 156 patients to report on the full cohort of 225 patients (i.e., chart reviews for the remaining 69 patients). The collected data is novel with respect to the literature, as we have collected three surveys for a variety of head and neck patients (i.e., not restricted to a subsite such as oropharynx). Possible study questions may include: QoL instrument comparisons, stratification of QoL scores based on site or other chart parameters, and ways in which the three surveys can clarify symptom constellations.

A second study looks assesses whether planned doses or delivered doses are more strongly associated with QoL scores. For this, we are using artificial intelligence to produce models to predict QoL scores. Our primary question is whether adaptive radiation therapy has the potential to improve patient-reported quality of life.

Please find below a collection of preliminary results and research notes.

Table 1: Cohort demographic and clinical characteristics

|  |  |
| --- | --- |
| Parameter | Full Cohort (n = 156) |
| Age in years, mean (SD) | 57.5 (10.9) |
| Gender, number (%)  Male  Female | 132 (84.6%)  24 (15.4%) |
| Initial BMI, mean (SD) | 28.2 (5.6) |
| ECOG, median (range) | 1 (1-3) |
| Charlson Comorbidity Index, median (range) | 4 (2-8) |
| Alcohol use, number (%)  Never  Former  Current – Light (males 0-15 drinks/week, females 0-10 drinks/week)  Current – Heavy (males >15 drinks/week, females >10 drinks/week) | 36 (23.1%)  12 (7.7%)  84 (53.8%)  24 (15.4%) |
| Tobacco use, number (%)  Never  Cumulative – Light (0-20 pack-years)  Cumulative – Heavy (>20 pack-years) | 63 (40.4%)  44 (28.2%)  49 (31.4%) |
| Primary tumor location, number (%)  Larynx  Hypopharynx  Oral Cavity  Oropharynx  Nasal Cavity  Nasopharynx  Unknown | 7 (4.5%)  3 (1.9%)  3 (1.9%)  99 (63.5%)  7 (4.5%)  26 (16.7%)  11 (7.0%) |
| T stage, number (%)  T0 – T2  T3 – T4  Tx | 72 (46.2%)  73 (46.8%)  11 (7.0%) |
| N stage, number (%)  N0  N1  N2  N3  NX | 24 (15.4%)  34 (21.8%)  83 (53.2%)  14 (9.0%)  1 (0.6%) |
| p16 status, number (%)  Negative  Positive  Unknown | 21 (13.5%)  101 (64.7%)  34 (21.8%) |
| Radiotherapy treatment, number (%)  Unilateral  Bilateral |  |
| Chemotherapy agent, number (%)  Carboplatin  Cetuximab  Cisplatin (Cisplatinum)  None | 3 (1.9%)  13 (8.3%)  128 (82.1%)  12 (7.7%) |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Question | Mean (SD) | Mild (0-4) | Moderate (5-6) | Severe (7-10) |
| 10. Your having a **dry mouth** at its WORST? | 4.3 (3.2) | 88 | 26 | 42 |
| 20. Your problem with **tasting food** at its WORST? | 3.1 (3.0) | 114 | 14 | 28 |
| 14. Your problem with **mucus** in your mouth and throat at its WORST? | 2.7 (3.1) | 116 | 17 | 23 |
| 15. Your difficulty **swallowing/chewing** at its WORST? | 2.6 (3.1) | 118 | 16 | 22 |
| 2. Your **fatigue (tiredness)** at its WORST? | 2.4 (2.8) | 122 | 16 | 18 |

Table 2: Top 5 MDASI items rated most severe (/10) according to cohort average and frequency of mild, moderate and severe ratings (/156)

Across all MDASI questions, 6 patients reported no symptoms (all items = 0), 56 patients rated symptoms no worse than mild, 28 patients rated symptoms no worse than moderate, and 66 patients reported at least one severe symptom. Note: highest-rated items are driven by the number of “severe” ratings.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Question | Mean (SD) | Mild (0-3) | Moderate (4) | Severe (5) |
| 10. It takes me longer to eat because of my swallowing problem. | 3.2 (1.4) | 70 | 55 | 30 |
| 5. I feel self-conscious when I eat. (negation asked) | 3.0 (1.5) | 83 | 40 | 32 |
| 16. I limit my food intake because of my swallowing difficulty. | 2.5 (1.4) | 103 | 36 | 15 |
| 11. People ask me, “Why can’t you eat that?” | 2.5 (1.3) | 106 | 41 | 8 |
| 7. Swallowing takes great effort. | 2.4 (1.2) | 116 | 29 | 10 |

Table 3: Top 5 MDADI items rated most severe (/5) according to cohort average and frequency of mild, moderate and severe ratings (/156)

Across all MDADI questions, 3 patients reported no symptoms (all items = 0), 11 patients rated symptoms no worse than mild, 64 patients rated symptoms no worse than moderate, and 77 patients reported at least one severe symptom.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Question | Mean (SD) | Mild (0-4) | Moderate (5-6) | Severe (7-10) |
| 7. Rate the frequency of sipping liquids to aid swallowing food. | 5.1 (3.4) | 66 | 29 | 58 |
| 8. Rate the frequency of sipping liquids for oral comfort when not eating. | 4.1 (3.2) | 89 | 24 | 40 |
| 5. Rate your mouth or throat dryness when eating food. | 3.5 (3.0) | 96 | 25 | 32 |
| 3. Rate your difficulty in swallowing solid food due to dryness. | 3.4 (3.1) | 103 | 20 | 30 |
| 6. Rate your mouth or throat dryness while not eating. | 3.2 (2.8) | 105 | 27 | 21 |

Table 4: Top 5 XQ ites rated most severe (/10) according to cohort average

Across all XQ questions, 9 patients reported no symptoms (all items = 0), 47 patients rated symptoms no worse than mild, 23 patients rated symptoms no worse than moderate, and 73 patients reported at least one severe symptom. Note: highest-rated items are driven by the number of “severe” ratings.

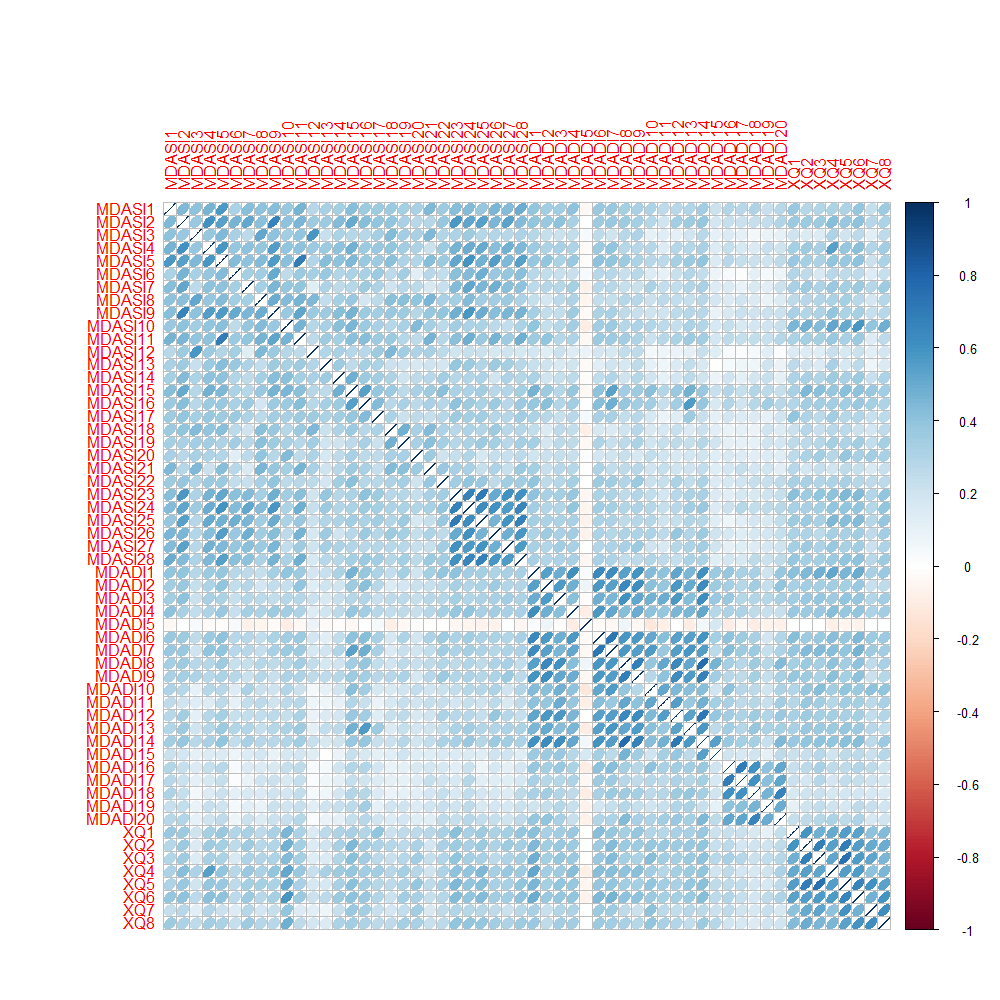
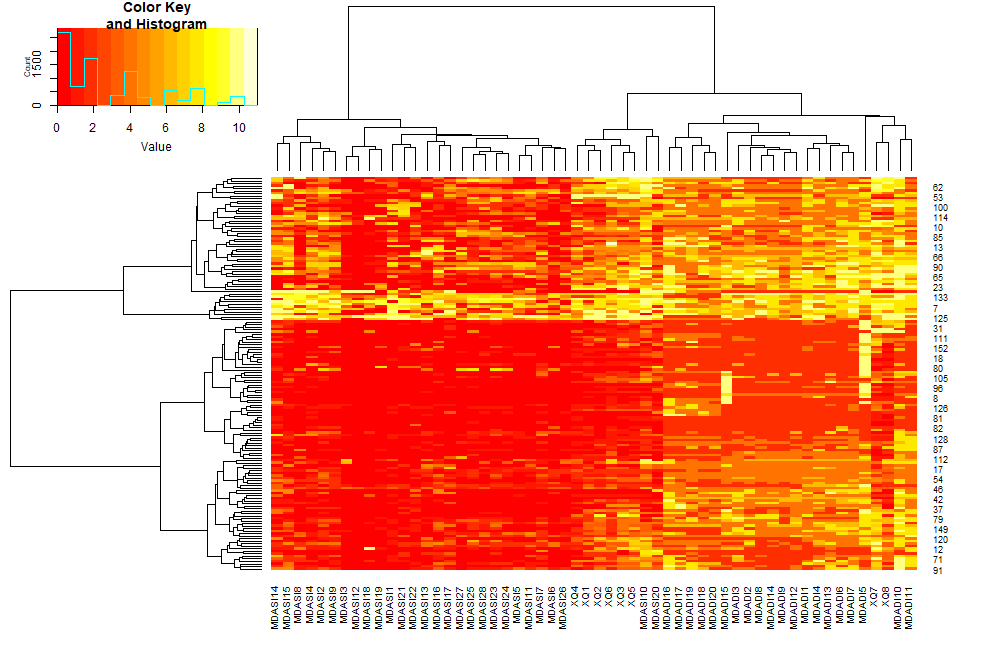


Figure 1: Correlations among QoL questions calculated using Kendall’s tau statistic. Within each survey, strong correlations exist among MDASI interference questions (MDASI23 – MDASI28), MDADI questions 1-14 (interesting as questions from various domains are interspersed throughout the survey), and all XQ questions. For the MDASI survey, poorer scores on general well-being questions (MDASI1-12) may correspond to a greater interference of symptoms on daily life (MDASI23-28). MDADI questions were moderately correlated with MDASI15-16 (difficulty swallowing/chewing and choking/coughing, respectively). XQ questions were moderately correlated with MDASI10 (having a dry mouth).

A close up of a map

Description automatically generated

Figure 2: Results of PAM-k and PCA when inputting all QoL questions. k-Medoid clustering identified two distinct groups of patients. Responses for each of the surveys were tightly grouped. XQ questions and corresponding responses appear intermediate to MDASI and MDADI.



**3**

**2**

**1**

**D**

**C**

**B**

**A**

Figure 3: Heatmap of survey responses with hierarchical clustering of patients (rows) and questions (columns). Clustering indicated the presence of four subgroups of patients with symptoms that were: mild (C), moderate(D), severe (A), and severe across all domains (B). Survey questions were generally grouped according to: (1) general symptoms and well-being, (2) xerostomia, and (3) dysphagia.

Combining the results of the patient and question clustering, we observe that: severe patients (A) primarily reported dry-mouth and swallowing symptoms with some general symptoms. Mild (C) and moderate patients (D) reported very few general symptoms, with a tendency to report poorer swallowing than dry-mouth symptoms. However, results may be influenced by the conversion of the 5-point MDADI scores to scores out of 10.

Additional notes: The three QoL question clusters show the three surveys are capturing different facets of a patient’s well-being. Also, similar questions specific to a given symptom will be given more weight during unsupervised analyses. It is possible that similar groupings could be identified by increasing the weighting on dry-mouth and swallowing fields in the MDASI, and omitting the MDADI and XQ surveys.

Based on hierarchical clustering, 55 patients (35.2%) had mild symptoms (cluster C), 44 patients (28.2%) had moderate symptoms (cluster D), 46 patients (29.5%) had severe symptoms (cluster A), and 11 patients (7.1%) had severe symptoms across all domains (cluster B). This is similar to the results of Eraj et al., with “65% ranging from none to moderate symptom burden, 35% with moderate-severe ratings for a subset of classically RT-related symptoms, and 2 of 79 patients (2.5%) with severe ratings of most items.” Our numbers are: 63.4%, 29.5% and 7.1%.

Running statistical tests on the four groups (i.e., which input variable is significantly different among the four groups) and mild/moderate vs. severe/very severe indicated that ECOG, initial BMI, final BMI, and gender as approaching significance. However, no variables met p < 0.05 criteria for statistical significance for the hierarchical clustering-based severity classifications.

General observations: patients in severe/very severe groups appear to have less setup variation, are younger, smaller initial and final BMI, poorer ECOG, have had treatment more recently; more replans; more females, more smoking, more PEG/NG tubes. This approximately corresponds with the behaviour we can see with PCA (below).

A picture containing text, map, table, photo

Description automatically generated

Figure 4: Principal components analysis of all numerical input parameters with hierarchical clustering symptom groups superimposed. Overlap of clusters is consistent with the fact that there are no statistically significant differences among clusters. However, mild/moderate clusters have greater dimension 2 values, the severe cluster has greater planned and delivered doses. The very severe cluster has smaller during-treatment systematic anatomical changes (back-projection of loadings) and poorer ECOG scores; potentially they may have had more recent treatment (back-projection of MonthsSinceTx)

Beyond the hierarchical clustering-based mild/moderate/severe/very severe ratings, we can consider statistically significant differences related to individual responses to questions. We can also consider the most severe quantile of rankings according to aggregate scores on QoL subdomains (i.e., physical, function, emotional).

Continuous independent variables were examined using Kruskal-Wallis test by ranks; categorical independent variables were examined using Fisher’s Exact test. Benjamini-Hochberg multiple testing corrections controlled the false discovery rate. P < 0.05 was considered statistically significant.

Table 5: Variables found to have statistically significant differences between mild/moderate and severe ratings for MDASI questions

|  |  |
| --- | --- |
| ***QoL Response*** | ***Statistically Significant Variable(s)*** |
| MDASI 3: Your nausea at its worst? (4) | Months since treatment (severe = longer since treatment) |
| MDASI 8: Your problem with lack of appetite at its worst? (13) | ECOG (severe = poorer ECOG) |
| MDASI 9: Your feeling drowsy (sleepy) at its worst? (14) | ECOG (severe = poorer ECOG) |
| MDASI 12: Your vomiting at its worst? (4) | Age (severe = younger age) |
| MDASI 21: Your mouth/throat sores at their worst? (14) | Gender (severe = female); p16 status (severe = p16 negative or unknown) |
| MDASI 24: Mood? (7) | Age (severe = younger and older age); p16 status (severe = negative or unknown) |

Table 6: Variables found to have statistically significant differences between mild/moderate and severe ratings for MDADI questions and aggregate scores

|  |  |
| --- | --- |
| ***QoL Response*** | ***Statistically Significant Variable(s)*** |
| MDADI 2: I am embarrassed by my eating habits (4) | Months since treatment (severe = longer since treatment) |
| MDADI 7: Swallowing takes great effort (10) | Alcohol intake (severe = less heavy drinking, more non- or previous- alcohol use) |
| MDADI 11: People ask me, why can’t you eat that? (8) | Alcohol intake (severe = less heavy drinking, more non- or previous- alcohol use) |
| MDADI 12: Other people are irritated by my eating problem (2) | Change in BMI (severe = more stable BMI); weight loss through the neck and shoulders (severe = less weight loss) |
| MDADI 13: I cough when I try to drink liquids (2) | Change in low-dose CTV volume (severe = greater decrease in low-dose CTV volume) |
| MDADI 19: I feel that I am swallowing a huge amount of food (3) | Age (severe = younger and older age) |
| MDADI\_E\_SUM, MDADI\_E\_AVE | T stage (severe = T3/T4), PEG/NG (severe = PEG/NG use) |
| MDADI\_F\_SUM, MDADI\_F\_AVE | Gender (severe = female), PEG/NG (severe = PEG/NG use) |
| MDADI\_TOTAL\_SUM, MDADI\_TOTAL\_AVE | PEG/NG (severe = PEG/NG use) |

Table 6: Variables found to have statistically significant differences between mild/moderate and severe ratings for XQ questions

|  |  |
| --- | --- |
| ***QoL Response*** | ***Statistically Significant Variable(s)*** |
| XQ1: Difficulty in talking due to dryness (17) | ECOG (severe = poorer ECOG), cancer site (severe = nasopharynx), p16 status (severe = p16 negative or unknown); alcohol intake (severe = less heavy drinking, more non-use) |
| XQ2: Difficulty in chewing due to dryness (21) | ECOG (severe = poorer ECOG) |
| XQ3: Difficulty in swallowing solid food due to dryness (30) | ECOG (severe = poorer ECOG) |
| XQ6: Mouth or throat dryness while not eating (20) | alcohol intake (severe = less heavy drinking, more non-use) |

Notes: ECOG and alcohol intake is significantly associated with many MDASI, MDADI and/or XQ responses. However, severe (rated = 5) MDADI scores only occur in a small subgroup of the cohort (e.g., as few as 2-3 patients of the cohort of 156). It appears that outcomes reported by patients with stable BMI throughout treatment and PEG/NG tube use and younger/older age may be contributing the most strongly to these results. Only MDADI aggregate scores corresponded to statistically significant differences in input variables.

Building on previous observations for QoL predictions:

The most straightforward approach to the QoL analysis seems to be predicting QoL responses using all other data (chart, planning CT, treatment plan, during treatment anatomical changes captured on CBCT, dose changes). However, we seem to be running into challenges in how the endpoints are defined. The MDADI scoring system is different than MDASI. MDASI and XQ share a similar 10-point Likert format, although the mild (1-4), moderate (5-6), severe (7-10) MDASI designations have not been validated for XQ. Isolating “severe” responses can result in very unbalanced datasets, where the experiences of as few as 2 patients can drive statistically significant differences.

Based on our COMP abstract, we may have more luck rolling in the ART focus by looking at planning objective violations. (Table below).

Thoughts: We were able to identify statistically significant differences in QoL scores when using raw numerical responses. Statistical significance was revealed when dichotomizing planning objective violations as factor variables.

Results may be most relevant by looking at numerical scores on QoL subdomains, supplemented by additional formatting of planned and delivered doses. Variables may be removed based on the number of missing values and indication of DIR noise in DIR vs. RO contour reviews (e.g., removing submandibular glands), and surrogates for DIR noise (e.g., change in brainstem volume and spinal cord volume).

Once this new spreadsheet is formatted, there are a few options for next steps:

One is to develop models to predict QoL subdomain scores. Given the restrictions on sample size, we can consider lasso linear regression and random forests, among others. Considering the predictive accuracy of models developed using planned dose vs. delivered dose. How predictors and predictive capability changes with QoL subdomain.

Perhaps this is too similar to the previous ART paper, however, although the scope would be different (can ART affect QoL), diverse focus based on QoL. Although study design is somewhat similar. Predictive model + categories + many endpoints + heuristic. It may be novel to look at how the predictive models compare. (Paired study?)

Another possibility is to look at descriptive approaches and unsupervised methods. “Swiss cheese” variable overlays of QoL clusters. Using AI to visualize the data problem.

Practically, we can also consider how this would fit into a clinical trial. What types of QoL improvements could we expect to see with ART? How can we refine PRO surveys to balance patient time with relevant questions to monitor QoL improvements in ART. Using the multi-endpoint patient-selection criteria, what QoL benefit could we expect? What refinements can we make to improve QoL impact? What questions should we include in a survey? All 3 surveys? Select subdomains? Streamlining the time commitment from 15 minutes to ensure that longitudinal studies are not too taxing.

Ways of formatting dose violations. Look at how raw planning objective violations correspond to QoL. Look at how ART/dose formatting correspond to QoL (Obj with trend analysis).

Table 7: Statistically significant differences in patient reported outcomes resulting from inter-fractional dose changes, anatomical changes, and patient characteristics

|  |  |  |  |
| --- | --- | --- | --- |
| ***Dose/Anatomical Parameter, Patient Characteristic*** |  | ***MD Anderson Dysphagia Inventory (MDADI)*** | ***Xerostomia Questionnaire***  ***(XQ)*** |
| Contralateral Parotid Gland Dmean  (26 Gy vs. 26 Gy) | Planned | - | - |
| Delivered | *People ask me, “Why can’t you eat that?”*  (2.1/5 vs. 2.9/5; p<0.01) | - |
| Brainstem D0.03cc  (54 Gy vs. 54 Gy) | Planned | *My swallowing difficulty has caused me to lose income.*  (1.6/5 vs. 3.3/5; p < 0.01) | - |
| Delivered | *My swallowing difficulty has caused me to lose income.*  (1.6/5 vs. 2.8/5; p < 0.01) | *Difficulty in talking due to dryness.*  (2.3/10 vs. 5.3/ 10; p = 0.03)  *Frequency of sleeping problems due to dryness.*  (2.2/10 vs. 5/10; p = 0.05) |
| Change in Low-Dose CTV Volume  (volume decrease  vs. increase) | N/A | - | *Difficulty in chewing due to dryness.* (3.4/10 vs. 4.1/10; p = 0.02)  *Mouth or throat dryness when eating food.* (3.1/10 vs. 5.0/10; p = 0.02) |
| Cancer Site  (other sites vs. nasopharynx) | N/A | *My swallowing ability limits my day-to-day activities.*  (2.1/5 vs. 2.9/5; p = 0.02)  *My swallowing difficulty has caused me to lose income.*  (1.6/5 vs. 2.2/5; p = 0.01) | *Difficulty in talking* (2.1/10 vs. 4.3/10; p < 0.01)*, chewing* (2.5/10 vs. 4.0/10; p = 0.04)*, and swallowing solid food* (3.2/10 vs. 4.8/10; p = 0.04) *due to dryness.*  *Frequency of your sleeping problems due to dryness.*  (2.1/10 vs. 3.8/10; p = 0.03)  *Mouth or throat dryness when not eating.* (3.0/10 vs. 4.5/10; p = 0.04) |

Mean differences in QoL scores are included above. Higher scores correspond to poorer patient-reported quality of life.

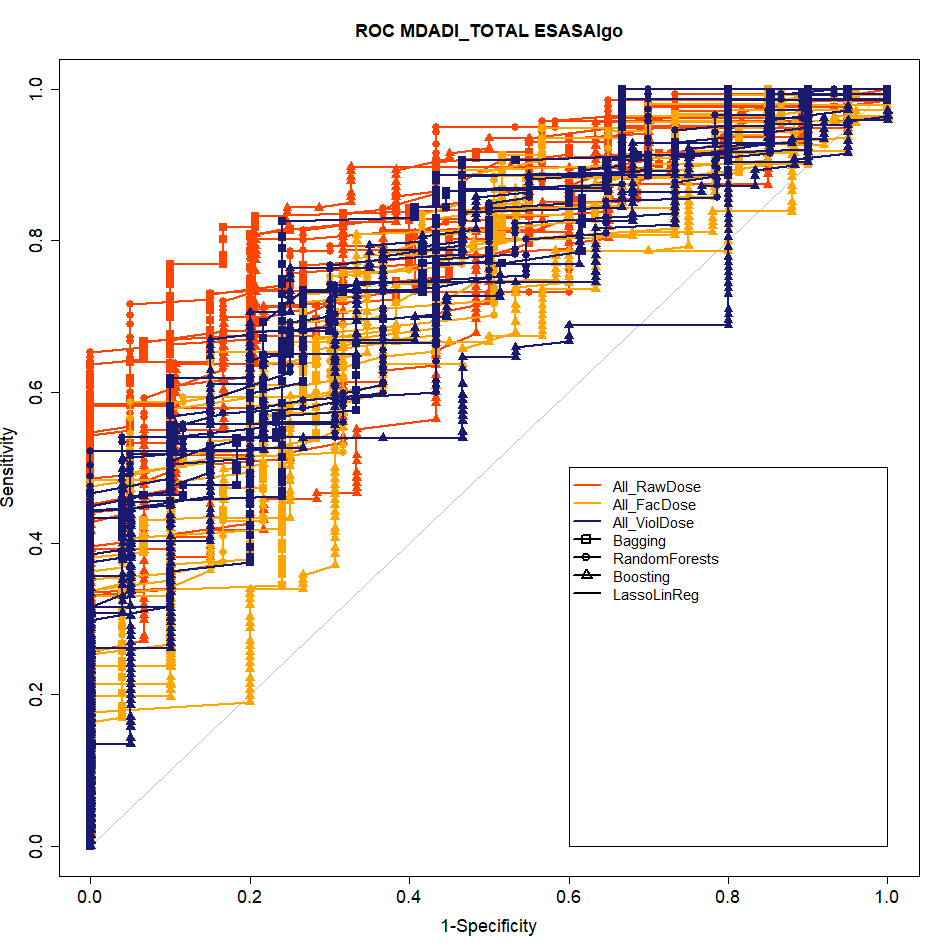
Modifying the dose/ART predictive modelling code for QoL. As with the dose/ART study, how well we are able to predict an endpoint is highly dependent on how the endpoint is formatted. So far for the QoL study, we have considered: 1) the maximum (worst) response for all questions assigned to a given QoL subdomain; 2) the quantile of patients with the worst average/total scores across a QoL subdomain; 3) a combined score based on the ESAS scoring algorithm (“Severe” if at least one item is >= 7, or if at least half of the responses are rated 5 or 6). There is perhaps some promise of predicting QoL scores, as indicated below.

Figure 6: ROC curves produced using 5-fold cross-validation with bagged/boosted trees and random forests. Pre-treatment data appears to best predict MDADI\_TOTAL scores according to the ESAS-motivated endpoint scoring. All\_RawDose uses planned dose to predict the endpoint; All\_FacDose uses delivered dose; All\_ViolDose uses all data.

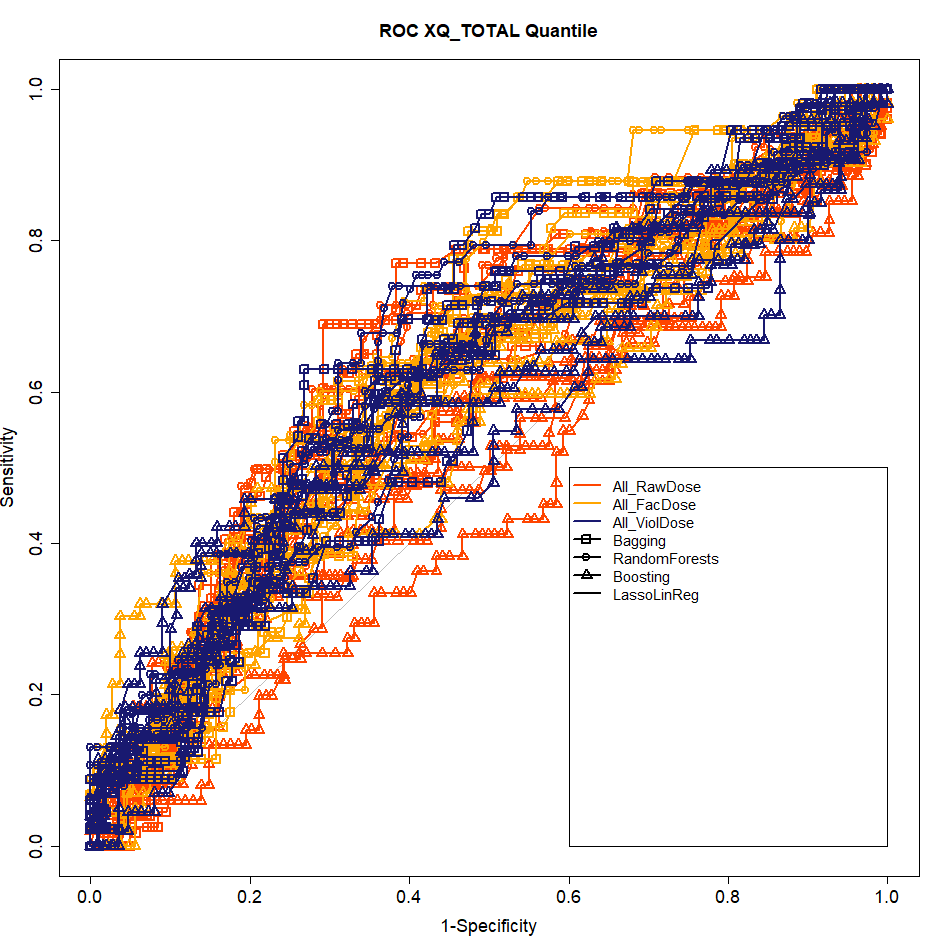
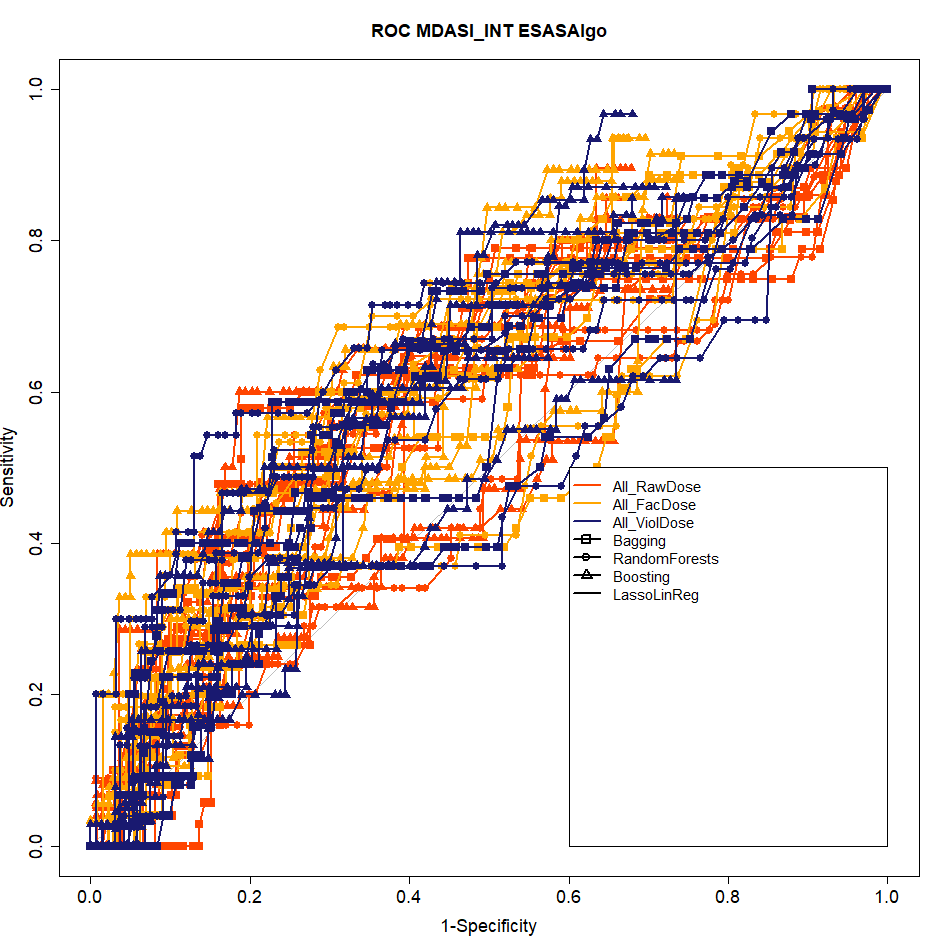


Figure 7: ROC curves produced using 5-fold cross-validation with bagged/boosted trees and random forests. During-treatment data may slightly better predict MDADI\_TOTAL scores according to quantile-motivated endpoint scoring. All\_RawDose uses planned dose to predict the endpoint; All\_FacDose uses delivered dose; All\_ViolDose uses all data.

Figure 8: ROC curves produced using 5-fold cross-validation with bagged/boosted trees and random forests. Interestingly, the predictive models could best predict the interference items among all of the MDASI subdomains. All\_RawDose uses planned dose to predict the endpoint; All\_FacDose uses delivered dose; All\_ViolDose uses all data.

As much of the QoL&ML paper will rely on being able to show that QoL scores can be predicted, worthwhile next steps may include:

* Further refining the endpoints: testing alternate interpretations of the ESAS scoring algorithm; applying the hierarchical clustering to subdomains. K-Medoid clustering? For the MDASI-CORE items, look at the scores for each item as its own endpoint. Perhaps total subdomain scores for mucositis/xerostomia/teeth/tasting/constipation are confounding our ability to predict them.
* Varying the input sets to identify which data is most predictive of a given QoL subdomain. For example, is pre-treatment patient chart data most predictive of MDASI-interference scores. Orange ROC curves have more input datapoints, which may cause their performance to be worse.
* Incorporating tree-based models for classification. Looking into what criteria classification models are optimized too. Building regression models on numerical endpoints has worked well before and gives a lot of flexibility for optimizing sensitivity and specificity. However, this may be possible for classification models and avoid the need for paired factor/numerical endpoints, which will be more challenging to produce for clustered scores.
* Looking at how many patients we could expect a QoL improvement in as a result of planning. I.e., how large of a trial would we need to run to observe a positive effect?
* Include submandibular gland information again. Although these OAR are prone to DIR noise, they may reveal something important for XQ\_NOTEAT scores…

If we are able to produce strong subdomain ROC curves, this would tell us:

* That our dataset is large enough to make predictions about QoL
* Which variables were of highest importance in the models
* For which endpoints delivered dose was as important or more important than planned dose
* Provides us with working models to infer the benefit of adaptive replanning on QoL outcomes

If delivered dose is more predictive of QoL than planned dose:

* Indicates that during-treatment dose corrections may have an effect on QoL, we can then look at which patients would benefit from replanning, and where QoL gains could be expected.
* This can streamline patient QoL surveys for longitudinal studies as part of an institutional clinical ART trial

If planned dose is more predictive of QoL than delivered dose:

* It may indicate that noise in the dose estimates confound the changes in QoL. We can review the results of the contouring study
* Planned dose may be correlated with pre-treatment chart data, we can look for these associations among the predictor sets
* Different dose parameters may better reveal associations with QoL. Current dose parameters are chosen based on functional endpoints. Perhaps it is possible that different parameters are more closely related to QoL.
* We may be able to fall back on the results of the dose/ART study. It may be possible to investigate a transitive argument. For example, if: A) planned parotid gland dose predicts XQ scores and B) planned parotid gland dose predicts delivered parotid gland dose, we can look at the thresholds on planned parotid gland dose and if: C) delivered parotid gland dose is associated with XQ scores.

If we cannot predict the QoL endpoints well:

* In this case, there are likely to be variables contributing to QoL that aren’t included in the study. For example, these may include psychosocial effects.
* Repeating the study with patient baseline results may help to clarify QoL associations.
* Sample size may be too small.

Study scope:

1. Is delivered dose more strongly associated with QoL than planned dose?
2. For which endpoints is this true?
3. Correcting doses back to planned (is this valid based on the literature?) what would be expected QoL gains from replanning? In how many patients? How would we identify them?
4. What does this mean for longitudinal ART QoL surveys in the future? Question/subdomain selection, patient inclusion/exclusion criteria, etc.
5. How do dose endpoints relate to QoL endpoints? Is the “signal” strong enough to predict relevant during-treatment dose changes with pre-treatment data? Is planning objective criteria relevant for QoL?