Clinical Experience with Knowledge-Based Planning

Lindsey Olsen, M.S.
Washington University in St. Louis
Disclosures

• Speaker Agreement with Varian Medical Systems
• License Agreement with Varian Medical Systems
Learning Objectives

1. Highlight the motivation for knowledge-based planning
2. Describe the clinical indication for KBP
3. Emphasize the importance of proper KBP model training and validation
Learning Objectives

1. Highlight the motivation for knowledge-based planning

2. Describe the clinical indication for KBP

3. Emphasize the importance of proper KBP model training and validation
Motivation

• Why Knowledge-Based Planning?
  o Leverage prior clinical and planning experience
  o Minimize repetition
  o Decision support (clinical sanity check)
  o Improve quality, efficiency, and automation
Motivation

Do IMRT planning goals guarantee optimal plans?

<table>
<thead>
<tr>
<th>H&amp;N</th>
<th>Bilateral Neck Treatment</th>
<th>Ipsilateral Neck Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTV</td>
<td>95% of PTV &gt; 95% of Rx; Max dose &lt; 110% of Rx</td>
<td>95% of PTV &gt; 95% of Rx; Max dose &lt; 110% of Rx</td>
</tr>
<tr>
<td>Spinal Cord</td>
<td>Max dose 40 Gy</td>
<td>Max dose 40 Gy</td>
</tr>
<tr>
<td>Spinal Cord + Margin</td>
<td>Max dose 52 Gy; &lt; 1% (or 1 cc) exceeds 50 Gy</td>
<td>Max dose 52 Gy; &lt; 1% (or 1 cc) exceeds 50 Gy</td>
</tr>
<tr>
<td>Optic Nerves, Optic Chiasm</td>
<td>Max dose 54 Gy</td>
<td>Max dose 54 Gy</td>
</tr>
<tr>
<td>Brainstem</td>
<td>Max dose 54 Gy; &lt; 1% exceeds 60 Gy</td>
<td>Max dose 54 Gy; &lt; 1% exceeds 60 Gy</td>
</tr>
<tr>
<td>Brain</td>
<td>Max dose 60 Gy; &lt; 1% exceeds 65 Gy</td>
<td>Max dose 60 Gy; &lt; 1% exceeds 65 Gy</td>
</tr>
<tr>
<td>Retina</td>
<td>Max dose 50 Gy; &lt; 5% exceeds 45 Gy</td>
<td>Max dose 50 Gy; &lt; 5% exceeds 45 Gy</td>
</tr>
<tr>
<td>Larynx</td>
<td>As low as possible; mean dose &lt; 45 Gy</td>
<td>As low as possible; mean dose &lt; 45 Gy</td>
</tr>
<tr>
<td>Upper Esophagus</td>
<td>As low as possible; mean dose &lt; 26 Gy</td>
<td>As low as possible; mean dose &lt; 25 Gy</td>
</tr>
<tr>
<td>Parotid</td>
<td>As low as possible; mean dose &lt; 26 Gy</td>
<td>As low as possible; mean dose &lt; 25 Gy</td>
</tr>
<tr>
<td>Pharyngeal Constrictors</td>
<td>As low as possible; mean dose &lt; 39 Gy</td>
<td>As low as possible; mean dose &lt; 39 Gy</td>
</tr>
<tr>
<td>Submandibular</td>
<td>As low as possible; mean dose &lt; 35 Gy</td>
<td>As low as possible; mean dose &lt; 35 Gy</td>
</tr>
<tr>
<td>Oral Cavity</td>
<td>Max 70 Gy; &lt; 5% exceeds PTV Rx</td>
<td>Max 70 Gy; &lt; 5% exceeds PTV Rx</td>
</tr>
<tr>
<td>Mandible</td>
<td>Less than PTV Rx; &lt; 5% exceeds PTV Rx</td>
<td>Less than PTV Rx; &lt; 5% exceeds PTV Rx</td>
</tr>
<tr>
<td>Unspecified Tissue</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Motivation

Yellow = Larynx
Red = Esophagus
Green = Partid
Magenta = PTV70
Cyan = PTV56
Motivation

**Yellow** = Larynx
**Red** = Esophagus
**Green** = Partid
**Magenta** = PTV70
**Cyan** = PTV56
Motivation

Variation in external beam treatment plan quality: An inter-institutional study of planners and planning systems

Benjamin E. Nelms PhD, Greg Robinson CMD, Jay Markham CMD, Kyle Velasco CMD, Steve Boyd CMD, Sharath Narayan CMD, James Wheeler MD, PhD, Mark L. Sobczak MD

Conclusions

There is a large inter-planner variation in plan quality as defined by a quantitative PQM score that measures the ability of the planner to meet very specific plan objectives. Plan quality was not statistically different between different TPS or delivery techniques and was not correlated to metrics of plan complexity. Certification and education demographics, experience and confidence level of the planner were not good predictors of plan quality.
Motivation

$\delta$ (prior) = 0.28 ± 0.24
$\delta$ (after) = 0.12 ± 0.13

In general...

- Significance of Knowledge-Based Planning

Prior Experience
- Training Plan 1
- Training Plan 2
- Training Plan 3
- Training Plan 4

Knowledge-Based Model

Improves Plan:
- Quality
- Standardization
- Efficiency
- Automation

Prior Experience
- Training Plan...

Barnes-Jewish Hospital • Washington University School of Medicine • National Cancer Institute • National Comprehensive Cancer Network
In general...

- **Significance of Knowledge-Based Planning**

Prior Experience
- Training Plan 1
- Training Plan 2
- Training Plan 3
- Training Plan 4
- Training Plan...

Knowledge-Based Model

Implements Plan:
- Quality
- Standardization
- Efficiency
- Automation

Does KBP deliver these claims?
In an inter-intuitional study it has been shown that the large inter-planner variation in plan quality

0%  a. depends on the planner’s experience
0%  b. is a direct result of the TPS
0%  c. is independent of planner’s experience
0%  d. depends on planner’s certification level
0%  e. is a direct result of the technique used
In an inter-institutional study it has been shown that the large inter-planner variation in plan quality depends on the planner’s experience, is a direct result of the TPS, is independent of planner’s experience, depends on planner’s certification level, or is a direct result of the technique used.

1. a. depends on the planner’s experience
2. b. is a direct result of the TPS
3. c. is independent of planner’s experience
4. d. depends on planner’s certification level
5. e. is a direct result of the technique used

Answer: c. is independent of planner’s experience

Learning Objectives

1. Highlight the motivation for knowledge-based planning

2. Describe the clinical indication for KBP

3. Emphasize the importance of proper KBP model training and validation
Outside Clinic Study

RT datasets for 20 clinically treated prostate IMRT plans from an outside institution transferred to Wash U pDVH DICOM tool

Patient 1:
RT DOSE
RT STRUCT

Export DICOM RT files from outside institution TPS

Import files into Wash U pDVH DICOM Tool

Appenzoller L.M., et al. Predictive DVH models developed at a large institution impact clinically relevant DVH parameters in IMRT plans at an unrelated radiotherapy facility, Oral presentation AAPM 2013.
Outside Clinic Study

- Clinic specific pDVH model created using institution’s own data
- Similar plan quality demonstrated for all patients
- No indication for improvement of clinically treated plans
Outside Clinic Study

Comparison against validated Wash U prostate model showed large improvements possible for rectum DVHs and small improvements for bladder DVHs for all patients.
Outside Clinic Study

- Five worst patients identified by sum of residuals between clinical DVH and predicted DVH
- Quantify improvements in clinical rectum and bladder DVHs with knowledge of pDVHs by replanning five worst patients

Replan five patients using optimization objectives exported from pDVH tool

Import replan dose matrix into DICOM tool and compare to original DVHs and pDVHs
All five patients replanned showed similar results…

### Table 3. Average Reduction in V65 and V40 for Rectum and Bladder

<table>
<thead>
<tr>
<th>Organ</th>
<th>V65(orig)-V65(replan)</th>
<th>dV65</th>
<th>V40(orig)-V40(replan)</th>
<th>dV40</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rectum</td>
<td>4.8%±2.3%</td>
<td>0.9%±1.1%</td>
<td>17.9%±10.3%</td>
<td>0.7%±1.4%</td>
</tr>
<tr>
<td>Bladder</td>
<td>3.4%±2.1%</td>
<td>0.4%±0.5%</td>
<td>6.0%±2.8%</td>
<td>0.6%±0.9%</td>
</tr>
</tbody>
</table>

Appenzoller L.M., et. al. Predictive DVH models developed at a large institution impact clinically relevant DVH parameters in IMRT plans at an unrelated radiotherapy facility, Oral presentation AAPM 2013.
Objective: To assess the impact of DVH prediction (pDVH) models and a standardized planning technique on post-operative endometrial IMRT treatment plan quality.
Institutional Plan Quality Study

Objective: To assess the impact of DVH prediction (pDVH)* models and a standardized planning technique on post-operative endometrial IMRT treatment plan quality.

- 25 post-op endometrial patient training set
- Replan with standard beams, contours, objectives
- Train pDVH models: Bladder, Bowel, Sigmoid, Rectum
- Model prediction accuracy: Sum of residual (SR) analysis
- Plan quality improvement: V40(clinical) – V40(replan)
Institutional Plan Quality Study

Objective: To assess the impact of DVH prediction (pDVH)* models and a standardized planning technique on post-operative endometrial IMRT treatment plan quality.

25 post-op endometrial patient training set

Replan with standard beams, contours, objectives

Train pDVH models: Bladder, Bowel, Sigmoid, Rectum

Model prediction accuracy: Sum of residual (SR) analysis

Plan quality improvement: V40(clinical) – V40(replan)

5 post-op endometrial patient validation set

Replan with standard beams, contours, objectives

pDVH IMRT optimization objectives

Model prediction accuracy: $d_{V40}$ and $d_{mean}$

Plan quality improvement: V40(clinical) – V40(replan)

Results

- The impact of using pDVH models and a standard planning technique is demonstrated by plan quality improvement in the 5 patient validation cohort as seen by a reduction in V40 and mean dose for all OARs compared with the original clinical plan.

<table>
<thead>
<tr>
<th>OAR</th>
<th>25 Patient Training Cohort</th>
<th></th>
<th>5 Patient Validation Cohort</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SR</td>
<td>V40(orig)-V40(replan)</td>
<td>Mean(orig)-Mean(replan)</td>
<td>V40(orig)-V40(replan)</td>
</tr>
<tr>
<td></td>
<td>(%)</td>
<td>(%</td>
<td>Gy)</td>
<td>(%)</td>
</tr>
<tr>
<td>Bladder</td>
<td>0.006 ± 0.045</td>
<td>8.8 ± 7.9</td>
<td>2.5 ± 1.7</td>
<td>9.8 ± 5.1</td>
</tr>
<tr>
<td>Bowel</td>
<td>0.017 ± 0.023</td>
<td>2.7 ± 2.4</td>
<td>2.4 ± 1.6</td>
<td>2.1 ± 2.1</td>
</tr>
<tr>
<td>Rectum</td>
<td>-0.007 ± 0.048</td>
<td>8.3 ± 8.8</td>
<td>3.2 ± 2.4</td>
<td>9.3 ± 5.9</td>
</tr>
<tr>
<td>Sigmoid</td>
<td>-0.012 ± 0.056</td>
<td>12.3 ± 13.9</td>
<td>3.5 ± 2.8</td>
<td>9.1 ± 14.8</td>
</tr>
</tbody>
</table>
Plan Quality Study Results

- The impact of using pDVH models and a standard planning technique is demonstrated by plan quality improvement in the 5 patient validation cohort as seen by a reduction in V40 and mean dose for all OARs compared with the original clinical plan.

<table>
<thead>
<tr>
<th>OAR</th>
<th>( V_{40}^{\text{orig}} - V_{40}^{\text{replan}} )</th>
<th>( \text{Mean}<em>{\text{orig}} - \text{Mean}</em>{\text{replan}} )</th>
<th>( \delta_{V_{40}} )</th>
<th>( \delta_{\text{Mean}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bladder</td>
<td>9.8 ( \pm ) 5.1</td>
<td>2.3 ( \pm ) 1.5</td>
<td>0.6 ( \pm ) 5.2</td>
<td>0.5 ( \pm ) 0.9</td>
</tr>
<tr>
<td>Bowel</td>
<td>2.1 ( \pm ) 2.1</td>
<td>0.5 ( \pm ) 0.6</td>
<td>1.7 ( \pm ) 1.4</td>
<td>0.5 ( \pm ) 1.4</td>
</tr>
<tr>
<td>Rectum</td>
<td>9.3 ( \pm ) 5.9</td>
<td>2.7 ( \pm ) 3.4</td>
<td>1.8 ( \pm ) 3.3</td>
<td>0.6 ( \pm ) 1.1</td>
</tr>
<tr>
<td>Sigmoid</td>
<td>9.1 ( \pm ) 14.8</td>
<td>1.8 ( \pm ) 2.3</td>
<td>1.3 ( \pm ) 5.4</td>
<td>0.4 ( \pm ) 1.5</td>
</tr>
</tbody>
</table>

Ongoing prospective clinical trial at Wash U to assess impact on plan quality and efficiency.
Multi-Institutional Study

- Secondary analysis of RTOG 0126

Multi-Institutional Study Results

- Results suggest decreased risk based on NTCP models

What Have We Learned?

- Treatment plan quality variability is a problem.
  - At Washington University in St. Louis
  - At independent clinics
  - At many of the academic and independent clinics that enrolled patients on RTOG 0126

- Does KBP/auto-planning address these issues?
  - Improves ability to systematically achieve high quality plan
  - Improves efficiency of treatment plan generation
  - Necessary to benchmark models against other institutions
Several published studies have demonstrated that knowledge-based planning models:

- are helpful QC for structure delineation (0%)
- can aid in plan quality improvement (0%)
- should never by used by a rad onc (0%)
- should only be used by a physicist (0%)
- are IMRT optimization algorithms (0%)
Several published studies have demonstrated that knowledge-based planning models

a. are helpful QC for structure delineation
b. can aid in plan quality improvement
a. should never be used by a rad onc
b. should only be used by a physicist
c. are IMRT optimization algorithms

Answer: b. can aid in plan quality improvement

Learning Objectives

1. Highlight the motivation for knowledge-based planning

2. Describe the clinical indication for KBP

3. Emphasize the importance of proper KBP model training and validation
Model Training and Validation

- Importance of systematic KBP model training and validation process:

Quality of KBP Model = Quality of Plan Created with KBP Model
Training and Validation Process

- Patient selection
- Model training and evaluation
- Model validation
- Clinical use of model
Training and Validation Process

• Patient selection
• Model training and evaluation
• Model validation
• Clinical use of model
Patient Selection: Geometry

- PTV / OAR Geometry
  - Similar target shape
  - Similar target location
  - Similar relative position of OARs to PTV
- CCMB ex.

Courtesy of J. Alpuche
Patient Selection: Guidelines

- Similar Clinical Objectives
  - Same PTV coverage/OAR sparing criteria
- Similar Clinical Trade-Offs
  - Importance of PTV coverage / OAR sparing
- PTV prescription dose can vary
  - Estimated DVHs will be scaled as a percentage of Rx dose

<table>
<thead>
<tr>
<th>Tissue</th>
<th>Bilateral Neck Treatment</th>
<th>Ipsilateral Neck Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTV/Spinal Cord/Spinal Cord + Margin/Brainstem/Brain/Retina/Larynx/Upper Esophagus/Parotid/Pharyngeal Constrictors/Submandibular/Oral Cavity/Mandible/Unspecified Tissue</td>
<td>95% of PTV &gt; 95% of Rx; Max dose &lt; 110% of Rx Max dose 40 Gy Max dose 52 Gy; &lt; 1% (or 1 cc) exceeds 50 Gy Max dose 54 Gy Max dose 54 Gy; &lt; 1% exceeds 60 Gy Max dose 60 Gy; &lt; 1% exceeds 65 Gy Max dose 50 Gy; &lt; 5% exceeds 45 Gy As low as possible; mean dose &lt; 45 Gy As low as possible; mean dose &lt; 45 Gy As low as possible; mean dose &lt; 26 Gy As low as possible; V60 &lt; 60 Gy As low as possible; mean dose &lt; 39 Gy As low as possible; mean dose &lt; 35 Gy Max 70 Gy; &lt; 5% exceeds PTV Rx Less than PTV Rx; &lt; 5% exceeds PTV Rx</td>
<td>95% of PTV &gt; 95% of Rx; Max dose &lt; 110% of Rx Max dose 40 Gy Max dose 52 Gy; &lt; 1% (or 1 cc) exceeds 50 Gy Max dose 54 Gy Max dose 54 Gy; &lt; 1% exceeds 60 Gy Max dose 60 Gy; &lt; 1% exceeds 65 Gy Max dose 50 Gy; &lt; 5% exceeds 45 Gy As low as possible; mean Dose &lt;25 Gy As low as possible; mean dose &lt; 25 Gy As low as possible; mean dose &lt; 25 Gy As low as possible; mean dose &lt; 25 Gy As low as possible; mean dose &lt; 25 Gy As low as possible; mean dose &lt; 24 Gy (contralateral) As low as possible; V60 &lt; 45 Gy As low as possible; mean dose &lt; 24 Gy (contralateral) As low as possible; mean dose &lt; 24 Gy (contralateral) Max 70 Gy; &lt; 5% exceeds PTV Rx Less than PTV Rx; &lt; 5% exceeds PTV Rx</td>
</tr>
</tbody>
</table>
Patient Selection: Patient Numbers

- Number of training patients increases as the model complexity increases.
- Model validation process is used to ensure the number of training patients is sufficient.
Patient Selection: Plan Quality

- **Training set plan quality**
  - Output of KBP model directly correlated to input
  - Statistical noise present in KBP training set can impact model behavior
- **QA of training set**
  - Clinically approved, safe treatment
  - Consider iterative process in model training to obtain adequate model
Plan Quality Considerations

- Ex. Prostate and Node model: OAR = Rectum
  - Poor correlation between actual and estimated DVH principal components for model trained with 70 mixed quality treatment plans
Plan Quality Considerations

- Ex. Prostate and Node model: OAR = Rectum
  - Good correlation between actual and estimated DVH principal components for model trained with 48 good quality treatment plans
Training and Validation Process

- Patient selection
- Model training and evaluation
- Model validation
- Clinical use of model
Model Training and Evaluation

- Review the model statistical results
- Review the clinical vs. estimated DVHs
- Review model outliers
  - Geometric and dosimetric

**Note:** Will discuss model evaluation and validation in context of Varian RapidPlan™. Specific steps will differ depending on algorithm and implementation of KBP software.
Review Model Statistics

- Assess model over-fitting
- Assess predictive ability of the model

Estimation model statistics for structure Rectum:

Model goodness of fit
- DVH's principal components average fit 0.998977 out of 1.0
- GED's principal components average fit 0.999374 out of 1.0
- Regression model parameters' coefficient of determination 0.772767 out of 1.0
- Regression model's parameters average chi square 1.14286
- Whole estimation model's fit 0.771433 out of 1.0
- Whole estimation model's average MSE 17.4678

Model goodness of estimation
- Mean squared error between original and estimate 0.00237184

Statistics outside boundaries:
- Proportion of histogram bins outside boundaries 40.4178
- Mean of absolute deviation of bins outside boundaries -0.00248895
- Mean squared error of bins outside boundaries 0.000618992
- Standard deviation of the error of bins outside boundaries 0.0132559
- Mean of the error of bins outside boundaries 0.0118805

Model was successfully trained with 41 out of 41 plans.

Model training done
Review Clinical vs. Estimated DVHs

- Model properly identifies variation in training set DVHs
Review Clinical vs. Estimated DVHs

- Clinical DVH > estimate → Outlier
- Clinically relevant parameter
Identify and Remove Outliers

- **Dosimetric outlier**
  - Clinical DVH substantially differs from estimated DVH based on a clinically significant parameter

- **Geometric outlier**
  - PTV volume/shape substantially differs from the majority of the training set
  - Structure volume/shape substantially differs from the majority of the training set
  - Positional relationship between structure and PTV substantially differs from the majority of the training set
Steps to Improve Model Quality

1. Add patients to address over-fitting
2. Remove geometric outliers or add similar patients
3. Remove or re-plan dosimetric outliers
Steps to Improve Model Quality

1. Add patients to address over-fitting
2. Remove geometric or dosimetric outliers or add similar patients
3. Re-plan possible dosimetric outliers

Iterative process
In statistical modeling, an outlier is defined as:

0%  a. a data point explained by the statistical model
0%  b. a data analysis technique
0%  c. a data point distant from other observations
0%  d. an application that takes input and generates output
0%  e. a method of understanding messages in the data
In statistical modeling, an outlier is defined as...

a. a data point explained by the statistical model
b. a data analysis technique
c. a data point distant from other observations
d. an application that takes input and generates output
e. a method of understanding messages in the data

Answer: c. a data point distant from other observations

Reference: Boris Iglewicz and David Hoaglin (1993), ÒVolume 16: How to Detect and Handle OutliersÓ The ASQC Basic References in Quality Control: Statistical Techniques, Edward F. Mykytka, Ph.D., Editor.
Training and Validation Process

- Patient selection
- Model training and evaluation
- Model validation
- Clinical use of model
Validation Patients

- Independent from patients used to train model
- Represent the range of patient geometries, plan geometries, and plan prescriptions for which the model will be clinically used
- Good plan quality
  - PTV coverage
  - OAR sparing
Clinical vs. Estimated DVHs

- Review that clinically approved plan is within DVH estimation range
- If it is not, it is possible that plan can be improved
Create Validation Plan w/ Model
Objective Selection

- IMRT objective selection
  - Ensures clinically acceptable plan that achieves model estimate
  - Based on prior clinical experience
  - Priorities and objectives tuned during model validation
Assess Clinical Acceptability

- Review validation plans as per normal institution clinical standards
- Isodose distribution
- Clinical guidelines (scorecard)
- PTV coverage
- Hotspots
- Population-based OAR DVH cut-points
- Plan technical integrity
Training and Validation Process

- Patient selection
- Model training and evaluation
- Model validation
- Clinical use of model
Clinical Use of Model

• Do not venture far from your validation set

• Consider automation/standardized protocols
  – Beam arrangement
  – Contouring guidelines
  – Plan quality reports (scorecards)

• Develop guidelines for clinical use
  – When should I use the model?
  – When should I plan manually?
Final Thoughts

- Proper model training and validation is necessary for the clinical use of knowledge-based planning models
  - Possibility for systematic errors

- KBP is an exciting advancement, with potential for future development
  - Potential to improve quality, efficiency, and standardization

- Does not replace human/clinician judgment
Acknowledgements

• Sasa Mutic, Ph.D.
• Kevin Moore, Ph.D.
• James Kavanaugh, M.S.
• Jun Tan, Ph.D.
• Jorge Alpuche, Ph.D.
• Jackie Wu, Ph.D.
• Min Yang, C.M.D.
• Jeff Michalski, M.D.
Questions?

www.siteman.wustl.edu