Improving Quality of Care in Radiation Therapy using AI in Physics Plan and Chart Review

Samuel M.H. Luk, Ph.D.

July 13, 2022

SAM Therapy Educational Course – Artificial Intelligence for QA
Introduction
What is Physics Plan and Chart Review in Radiation Oncology?

• “Assure MUs are correct, all machine parameters used for patient setup are correct, additional setup instructions are correct, quality of the plan meets department standards, all signatures, prescriptions are recorded” – TG-40

• Initial plan review has shown to be the most effective individual QC check for detecting high severity incidents

Ford et al. IJROBP, 84(3): e263 - e269 (2021)
Recommendations on Initial Treatment Plan Review

TG-275 – Strategies for Effective Physics Plan and Chart Review in Radiation Therapy

• Use a risk-based approach (FMEA) to develop recommendations to physics plan and chart review

• Photon/Electron EBRT initial plan/chart review checks
  • Patient assessment
  • Simulation
  • Treatment planning
  • Data Transfer (for some combinations of TPS and OIS)
Recommendations on Initial Treatment Plan Review

MPPG 11a – Plan and Chart Review in External Beam Radiotherapy and Brachytherapy

• Goal: Provide recommendations on plan/chart review in the form of example lists of items to check for medical physicists and other clinical staff

• Initial EBRT Treatment Plan/Chart Review Items for Medical Physicists
  • Plan integrity check
    • E.g. Isocenter/initial reference point
    • Plan Quality and dose metrics reasonable
  • Preparation in RO-EMR
    • E.g. Prescription
    • Tolerance table
Components of Initial Treatment Plan Reviews

• Items require simple check
  • Examples:
    • Prescription matches order
    • Dose constraints are fulfilled
    • Data transfer accuracy

• Items require logical judgement
  • Examples
    • Prescription is suitable for tumor type
    • Treatment technique fits the patient anatomy
Automation and Tools to Support Initial Plan Review

• Multiple in-house software and commercial products are developing/developed to assist initial plan review

• Perform mostly rules-based checks
  • e.g. Rx matches, DVH constraints met etc.
  • Good for items require only simple checks

• They are great tools to improve efficiency and effectiveness as recommended in MPPG 11.a and TG 275
Rules-based Algorithms

- First order logic
  - E.g. If the isocenter of setup beams is different from treatment fields, then it is flagged as an error

- Advantages
  - Fast
  - Transparent
  - Good at finding static errors (protocols)

- Disadvantages
  - Difficult to check complex relationships
  - Need to update manually
Artificial Intelligence for Plan Review
AI as an Assistive Tool in Physics Plan Review

- Can factor in different information of a treatment plan to assist physicists on judging the appropriateness of the technical aspects of treatment
  - E.g. is the prescription appropriate, should a bolus be used etc.
- Can be kept up-to-date to latest clinical development by re-training the models with latest clinical data
Outlier Detection Model

• Outlier detection model using a k-mean clustering algorithm for plan review of prostate cases planned with ‘four-field’ box
• Look for outliers in MU as well as beam energy

Figure 3. Outlier detection rate as a function of error level for each of the MU features and all MU features combined. The error bar shows one standard deviation.
Bayesian Network-based algorithm

- **Artificial intelligence**
  - Mimic human reasoning to some degree by learning from data

- **Advantages**
  - Address points that require judgement
  - Leverage clinical data and adapts to local practice and update with latest practice
  - Interpretable

- **Disadvantages**
  - Slower running speed
  - Probabilistic results
Error Detection Bayesian Network (EDBN)

- EDBN was developed to help detect potential errors in treatment plans
- Provide assistant on judging the appropriateness of treatment parameters given the diagnostic parameters
- 4 categories of parameters
  - Diagnostic
  - Prescription
  - Plan and field parameter
  - Setup

Effectiveness of the Network

- Testing cases with manually embedded errors
- Types of errors
  - Prescription
  - Plan/Beam
  - Setup
- Area Under Curve = 0.89
Multi-Layered Approach using Rules and AI

- Combining the advantages of Bayesian Network and Rules

- Rules
  - Fast and good at identifying static errors

- Bayesian Network
  - Can mimic human logic and leverage clinical data to adapt local practice

<table>
<thead>
<tr>
<th></th>
<th>Bayes net</th>
<th>Rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Errors of judgment</td>
<td>✔️</td>
<td>✗</td>
</tr>
<tr>
<td>Maintenance/Updating</td>
<td>✔️</td>
<td>✗</td>
</tr>
<tr>
<td>Complex relationships</td>
<td>✔️</td>
<td>☑️</td>
</tr>
<tr>
<td>Transparency</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Speed</td>
<td>✗</td>
<td>✔️</td>
</tr>
<tr>
<td>Static Errors (protocols)</td>
<td>✗</td>
<td>✔️</td>
</tr>
</tbody>
</table>
Plan Check Tool - Rules

### Rule check Results tab

<table>
<thead>
<tr>
<th>Rule check</th>
<th>Result</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>beamDoseCheck</td>
<td>pass</td>
<td>check that sum of all beams doses add up to Rx they belong to</td>
</tr>
<tr>
<td>RxNotApproved</td>
<td>pass</td>
<td>verify that current prescription status is 'approved'</td>
</tr>
<tr>
<td>modalityMatch</td>
<td>pass</td>
<td>verify that prescription modality matches beam modality</td>
</tr>
<tr>
<td>BeamSetIsosCheck</td>
<td>pass</td>
<td>check that all beams in an Rx have same iso</td>
</tr>
<tr>
<td>CheckForDRRs</td>
<td>pass</td>
<td>check that all static fields have DRRs (including KV)</td>
</tr>
<tr>
<td>paceMakerCheck1</td>
<td>pass</td>
<td>warn use of 18MV for pacemaker pt</td>
</tr>
<tr>
<td>MUsegmentmax</td>
<td>pass</td>
<td>Check that MU per segment is not larger than 999</td>
</tr>
<tr>
<td>MUsegmentmin</td>
<td>pass</td>
<td>Check that MU per segment greater than 5 for non-VMAT plans</td>
</tr>
</tbody>
</table>

### Universal rules

- Select patient, prescription and site

<table>
<thead>
<tr>
<th>Rule check</th>
<th>Result</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>radicalDocExist</td>
<td>-not required-</td>
<td>verify that MU second check calc docs are in</td>
</tr>
<tr>
<td>QAExist</td>
<td>fail: no QA documented (audit)</td>
<td>verify that QA docs are in (VMAT, IMRT, SBRT)</td>
</tr>
<tr>
<td>respDocExist</td>
<td>-not required-</td>
<td>verify that phys response is in (TBI, SBRT)</td>
</tr>
<tr>
<td>SBRTRepartDoc</td>
<td>-not required-</td>
<td>verify that dosi report exists for SBRT</td>
</tr>
<tr>
<td>ISOsetupCheck</td>
<td>pass</td>
<td>check that setup beams have same iso as treatment fields</td>
</tr>
<tr>
<td>TBIsetupCheck</td>
<td>-not required-</td>
<td>check if most recent roadmap and CCP are approved and spec proc is in</td>
</tr>
</tbody>
</table>

### Site-specific rules + documentation
Site-specific networks are pre-built for the web app using local clinical data from Mosaiq.

Probability of each parameter in the network is calculated.
Bayesian Network – Web Application (Cont.)

"Probabilistic Results" tab

Alert!

Make sure it is correct!
Bayesian Network for Prescriptions

- Detect errors in physician orders/Rx
- Divided the prescription orders into 3 groups: single Rx, concurrent boost and sequential boost
- Detect errors in new orders given the disease information
Quality Assurance on Contours

Groupwise Conditional Random Forests for Automatic Shape Classification and Contour Quality Assessment in Radiotherapy Planning

Chris McIntosh*, Igor Svistoun, and Thomas G. Purdie

Automatic contouring QA method using a deep learning–based autocontouring system

Dong Joo Rhee1,2, Chidimma P. Anakwenze Akinfenwa3, Bastien Rigaud4,5, Anuja Jhingran6, Carlos E. Cardenas7, Lifei Zhang4, Surendra Prajapat8, Stephen F. Kry9, Kristy K. Brock10, Beth M. Beadle11, William Shaw12, Frederika O’Reilly13, Jeannette Parkes14, Hester Burger15, Nazia Fakie16, Chris Trauernicht17, Hannah Simonds18, Laurence E. Court19

Automatic detection of contouring errors using convolutional neural networks

Dong Joo Rhee2, The University of Texas Graduate School of Biomedical Sciences at Houston, Houston, TX 77030, USA, Department of Radiation Physics, Division of Radiation Oncology, The University of Texas MD Anderson Cancer Center Houston, TX 77030, USA
Carlos E. Cardenas2, Department of Radiation Physics, Division of Radiation Oncology, The University of Texas MD Anderson Cancer Center Houston, TX 77030, USA
Hesham Eltawil2, Department of Radiation Oncology, Division of Radiation Oncology, The University of Texas MD Anderson Cancer Center Houston, TX 77030, USA
Rachel McCarroll2, Department of Radiation Oncology, The University of Maryland Medical System, Baltimore, MD 21201, USA
Lifei Zhang4, Jinzhong Yang2, Department of Radiation Physics, Division of Radiation Oncology, The University of Texas MD Anderson Cancer Center Houston, TX 77030, USA
Adam S. Garden2

Investigating the potential of deep learning for patient-specific quality assurance of salivary gland contours using EORTC-1219-DAHANCA-29 clinical trial data

Hanne Nijhuis12, Ward van Rooij4, Vincent Gregoire5, Jens Overgaard6, Berend J. Slotman4, Wilko F. Verbakel3 and Max Dahele3

1Department of Radiation Oncology, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands; 2Department of Radiation Oncology, Centre Leon Berard, Lyon, France; 3Department of Clinical Medicine – Department of Experimental Clinical Oncology, Aarhus University, Aarhus N, Denmark

Nijhuis et.al. Acta Onco 1863463 (2020)
Rhee et.al. JACMP e13647 (2022)
Challenges of Development and Implementation on AI for Plan Review
What are the Hurdles?

• Standardization of data content, data format, data structure, and nomenclature

• Data Extraction

• Model generalizability and external validation

• Model Interpretability

• Quality assurance procedures for AI tools

• Simulated plans with errors for test and validation

• Trust on AI-generated results
Standardization, Data Extraction, Model Generalizability and Interpretability

- Collaboration between UVM, UW and Maastro
- Tested the network on cases with simulated errors in Maastro
- Multiple networks are trained (UW, Maastro, UW+Maastro)
- Performance has shown to be reduced

### Causes of Change in Performance

<table>
<thead>
<tr>
<th>Institution/Clinical Settings</th>
<th>Linacs</th>
<th>Treatment planning system</th>
<th>Oncology information system</th>
</tr>
</thead>
<tbody>
<tr>
<td>UW</td>
<td>Elekta</td>
<td>RayStation</td>
<td>Mosaiq</td>
</tr>
<tr>
<td>Mastro</td>
<td>Varian</td>
<td>Eclipse</td>
<td>Aria</td>
</tr>
<tr>
<td>UVMMC</td>
<td>Elekta</td>
<td>Pinnacle</td>
<td>Mosaiq</td>
</tr>
</tbody>
</table>

Table 1. Differences in technologies between the participating institutions.

<table>
<thead>
<tr>
<th>Type of error</th>
<th>Mean AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bolus</td>
<td>0.76</td>
</tr>
<tr>
<td>Collimator angle</td>
<td>0.70</td>
</tr>
<tr>
<td>Table angle</td>
<td>0.90</td>
</tr>
<tr>
<td>Prescription dose</td>
<td>0.55</td>
</tr>
<tr>
<td>Gantry angle</td>
<td>0.67</td>
</tr>
<tr>
<td>Overall</td>
<td>0.68</td>
</tr>
</tbody>
</table>

Table 2. AUCs for different types of errors in the external validation of UW-trained EDBN on Masstro data.
Improvements that We Are Working On

- Map the data of each clinic to a standardized list
- New network structure to accommodate all clinical profiles
- Distributed learning to adopt to individual clinical practice vs pooled data

Quality Assurance Procedures for AI Model

- Independent QA procedures of AI products are required
  - Performance of AI model will decay over time
- QA needs to ensure a consistent performance and require update of the model when it is under-performing
- No standards or guidelines yet for AI performance metrics

Luk et.al. Clinical Oncology 34(2):89-98 (2021)
Simulated Plans for Test and Validation

AAPM webpage ➔ Quality & Safety Resources ➔ Simulated Error Training for the Physics Plan Review

Credit: Perry Johnson and WGPE, AAPM 2022 MO-FG-201
<table>
<thead>
<tr>
<th>#</th>
<th>Error details</th>
<th>T1a # Failure mode</th>
<th>T1c # Plan/chart check</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Rectal balloon contrast not over-ridden to air</td>
<td>Wrong dose calculated due to contrast override</td>
<td>Density overrides applied as needed (ex. High-Z material, contrast, artifacts, etc.)</td>
</tr>
</tbody>
</table>

A rectal balloon with contrast is evident when reviewing the CT dataset. During treatment this will be filled with air according to the sim order. A density override is missing. In Raystation, this would be evident in the ROI Matl column (see arrow) and in the treatment plan report.
Trust on AI-Generated Results

- Participating physicists expressed difficulties to understand how to interpret results of probabilistic component generated from AI.
- Presentation and frequency of false positive results present a challenge of tradeoffs between trust, efficiency, and efficacy.
Summary

• Initial plan review is an important safety barrier in radiotherapy processes

• Despite its importance, AI development is not commonly found in plan review due to multiple challenges

• There are still a lot of opportunities to develop AI to assist medical physicists on plan review in conjunction with the automated rules-based tools
Thank You!