Large-Scale Multi-Patient Radiotherapy Data Mining Framework with Commercial Plan Reporting Tool

Mojtaba Moazzezi[1], Kevin Moore[1], Kurt Sysock[2], Xenia Ray[1] , Kelly Kisling[1]

[1]UC San Diego Health, [2]Radformation Co, New York, NY

UC San Diego Health

RETHINKING MEDICAL PHYSICS

INTRODUCTION

• The curation of sufficiently large treatment planning datasets is typically a labor-intensive process that impedes many studies in radiotherapy.

• As clinical radiotherapy investigations become more data-driven based on analysis of large cohorts of prior data, the efficient analysis of multi-patient samples becomes even more critical for machine learning applications, automated planning validation, and outcomes studies.

• **The purpose of this project was to develop a data extraction tool based on an Application Programming Interface (API) for ClearCheck™ that allows for the efficient batch processing of arbitrarily large numbers of datasets from the Eclipse™ treatment planning system.**

METHOD

• The ClearCheck API is an integrated commercial plan evaluation tool for Eclipse, where its classes and functionally can be called in other C# applications.

• We created a stand-alone executable script that uses the ClearCheck API to extract treatment planning data in batch, including the prescription, beam parameters, DVH, and structure names, and export them as Java script object notation (JSON) files.

• We then used Matlab to parse the JSON files and analyze the data for two separate studies:

(1) Classifying breast patients by treatment technique

(2) Comparing DVH data for planned versus adapted treatments of prostate patients
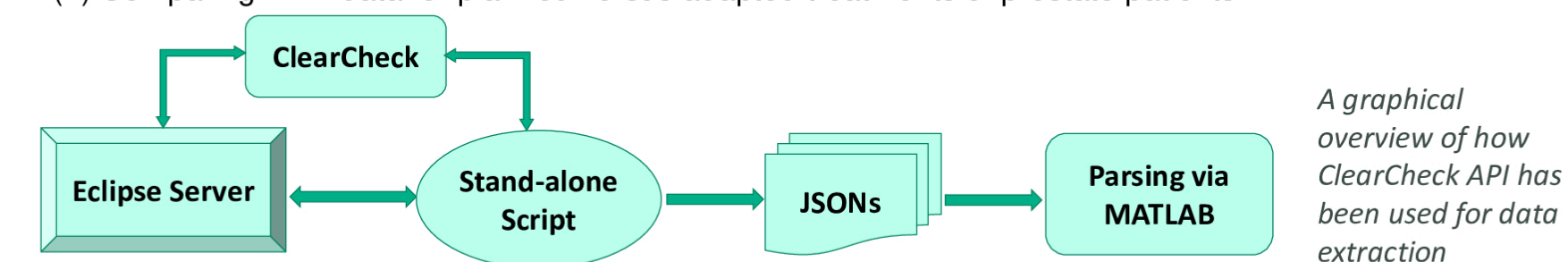
A graphical overview of how ClearCheck API has been used for data extraction
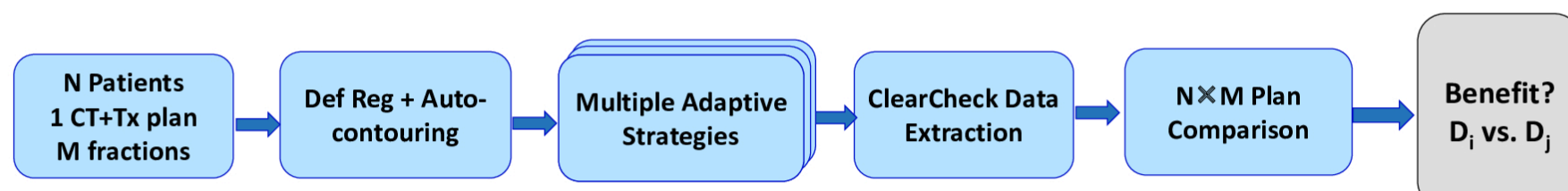

A graphical overview of multiple plan comparison workflow for prostate adaptive study

RESULTS

Using the ClearCheck API enabled us to:

• Mine the treatment planning data for 1800+ patients who had received radiotherapy for breast cancer from 2010-2018 in our clinic, and efficiently classify them based on treatment technique, laterality, and patient position in preparation for training a knowledge-based dose prediction model.

• Allowed us to easily extract the DVH data for 500+ plans (planned and adapted) across 25 prostate cancer patients, which facilitated the comparison of multiple adaptive treatment strategies (*results can be found in a separate poster: PO-GeP-M-404*).

1775 patients who received radiotherapy for breast has been classified based on applied treatment techniques (bilateral cases have been excluded).
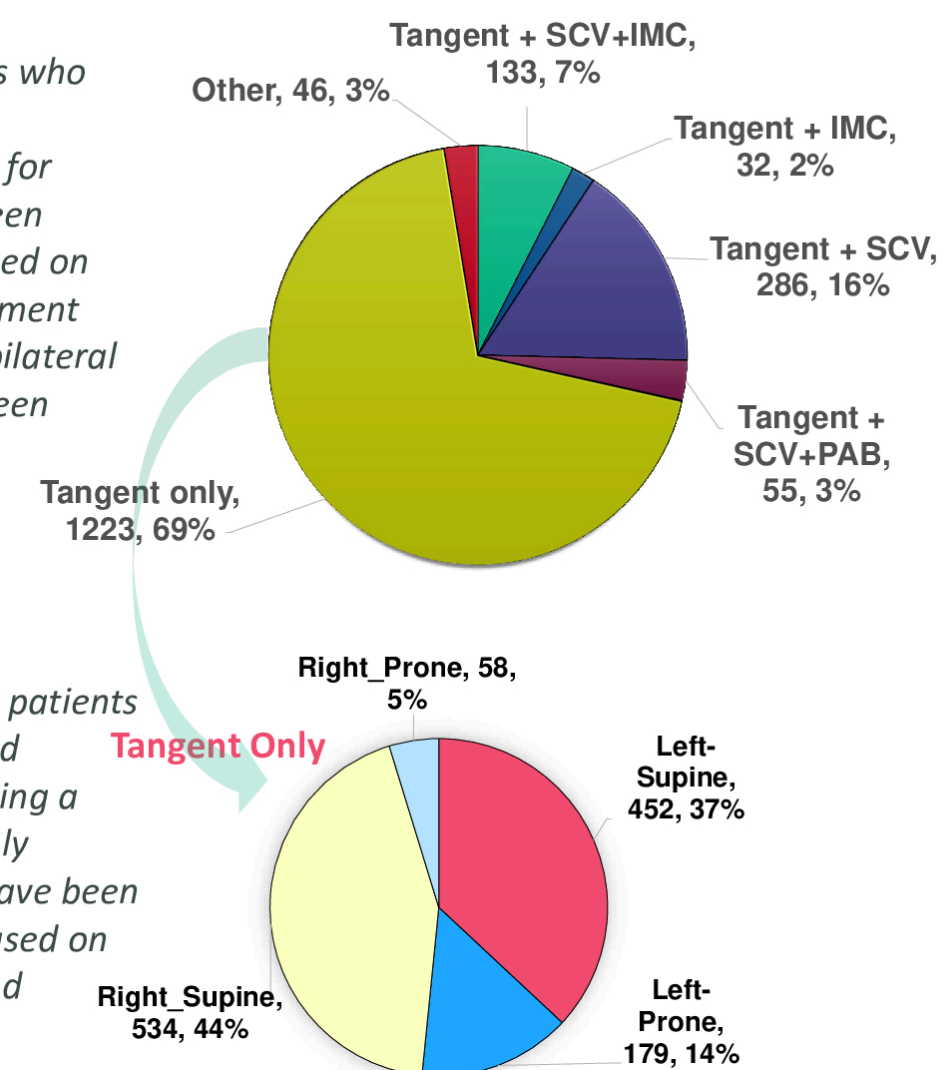
1223 breast patients who received radiation using a tangents-only technique have been classified based on laterality and orientation.

CONCLUSIONS

Using the ClearCheck API, we were able to readily extract and analyze vast amounts of treatment planning data for two separate studies requiring hundreds to thousands of datasets.

Tools such as the ClearCheck API can help researchers overcome the challenges of data curation and streamline the aggregation of datasets large enough to answer many open questions in radiotherapy.

ACKNOWLEDGEMENTS

CONTACT INFORMATION

Mojtaba Moazzezi, mmoazzezi@health.ucsd.edu