Automated Adaptive Decision Making with Deep Leaning Neural Network

Issam El Naqa, PhD, DABR, FAAPM, FIEEE
Chair of Machine Learning
Senior Member of Radiation Oncology

07/29/2020

MOFFITT CANCER CENTER

1

Acknowledgements

Moffitt
Yoga Balagurunathan, PhD
Tom Dilling, MD
Heiko Enderling, PhD
Jason Fleming, MD
Jennifer Frakes, MD
Robert Gillies, PhD
Luo Harrison, MD
Sarah Hoffe, MD
Peter Johnstone, MD
Kujtim Latifi, PhD
Eduardo Moros, PhD
Jose Penagaricano, MD
Dana Rollison, PhD
Ahmad Tarhini, MD
Michal Tomaszewski, PhD
Stu Wasserman, MS

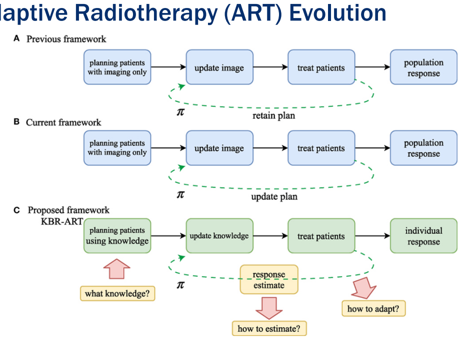
Umich
James Balter, PhD
Paul Carson, PhD
Roy Clarke, PhD
Kyle Cuneo, MD
Ivo Dinov, PhD
Judy Jin, PhD
Ted Lawrence, MD PhD
Dan McShan, PhD
Martha M. Matuszak, PhD
Jean Moran, PhD
Dipankar Ray, PhD
Al Rehemtulla, PhD
Ben Rosen, PhD
Randall K. Ten Haken, PhD
Xueding Wang, PhD

Trainees and staff
- Noora Ba Sunbul
- James Coates, PhD
- Sunan Cui, PhD
- Jamalina Jamaluddin
- Yi Luo, PhD
- *Dipesh Niraula, PhD*
- *Abe Oraiqat, PhD*
- Julia Pakela
- Wenbo Sun, PhD
- Huan-Hsin Tseng, PhD
- Lise Wei, PhD
- Wei Zhang, PhD

NIH National Institutes of Health

2

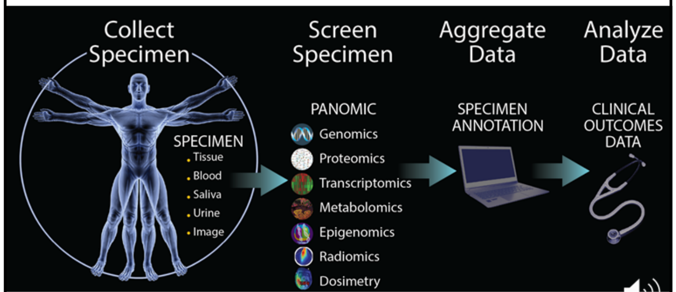
Adaptive Radiotherapy (ART) Evolution



Tseng, Frontier, 2018

3

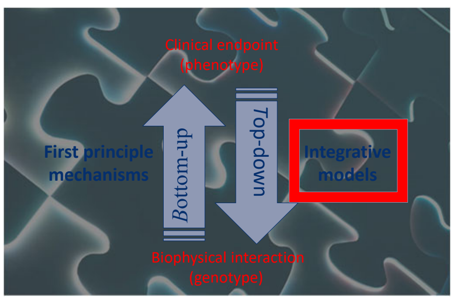
What? - ART in era of -omics



El Naqa et al., PMB (Journal highlights), 2017

4

Which? - radiotherapy response estimates

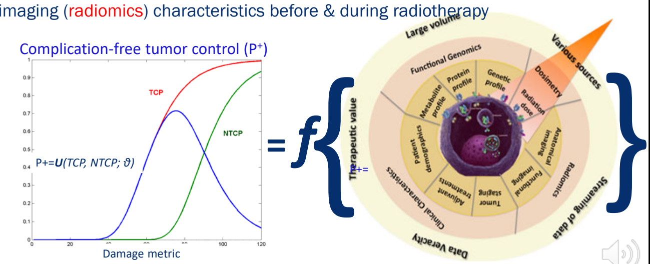


El Naqa, Wires, 2014

5

Integrative radiobiological modeling

Tumor control (TCP)/Normal tissue complication (NTCP) are multi-factorial and depend on: radiation dose and patients' genomic (radiogenomics) and imaging (radiomics) characteristics before & during radiotherapy

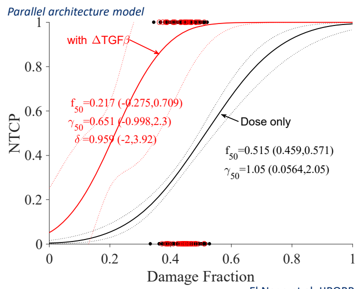


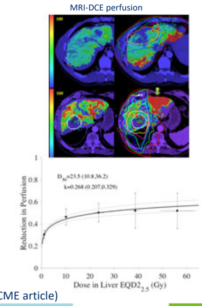
El Naqa, Methods, 2016

6

Analytical NTCP Liver Modeling: Dose+ cytokines + imaging

ALBI changes in liver cancer

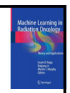




El Naqa et al, IJROBP, 2018 (CME article)

7

Outcome modeling by Machine learning (ML)



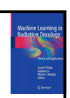
- *Generative* models
  - Model class-conditional PDFs and prior probabilities (Bayesian networks, Markov models)
    - *To predict you need to know the system*

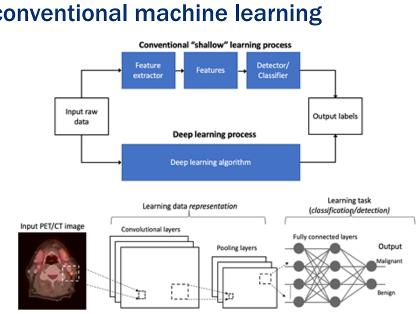
- *Discriminant* models
  - Directly estimate posterior probabilities (logistic regression, neural networks, CNN, random forests, SVM)
    - *Predict without knowing the system*

Tseng, Oncology, 2018

8

Deep vs conventional machine learning



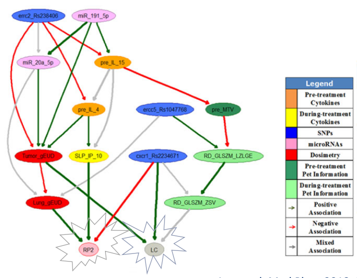


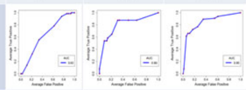
Zaidi & El Naqa, Annu. Rev. Biomed. Eng, 2021

9

Multi-Objective Generative (Bayesian) Models

A MO-BN can be used to predict multiple radiation outcomes simultaneously, which provides opportunities of finding appropriate treatment plans to solve the trade-off between competing risks.



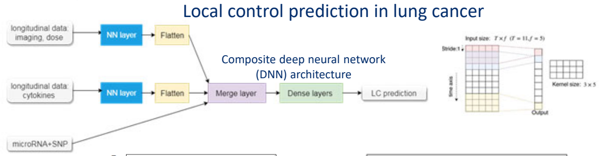
	Lung_gEUD and Tumor_gEUD	Pre-Treatment BN for joint prediction of LC and RP2	During-Treatment BN for joint prediction of LC and RP2
AU-FROC	0.63	0.8	0.85
95% CI	0.53-0.77	0.66-0.85	0.71-0.89
FROC			

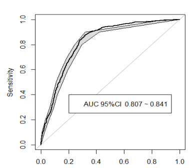
Luo et al, Med Phys, 2018 (Editor's Choice)

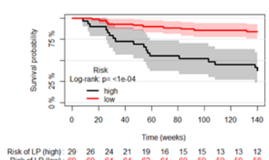
10

Radiogenomics modeling with deep ML

Local control prediction in lung cancer



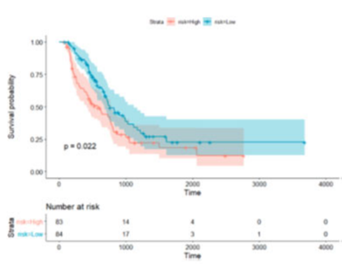




Cui et al, IEEE TRPMS, 2018 (Best of Physics, ASTRO)

11

Deep Survival Radiomics model for Liver Cancer I

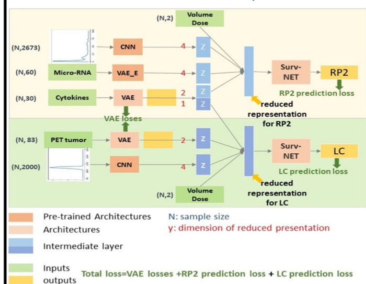


Wei et al, Physica Medica, 2021

12

Multi-objective radiogenomics model with deep survival neural networks I

DSNN prediction and comparison

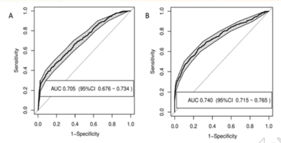


20 times of 5-fold cross validations

C-index (95%CI)	RP2	LC
NN-com	*0.705 (0.676~0.734)*	*0.740 (0.715 ~0.765)*
NN-DVH	0.660 (0.630~0.690)	0.727 (0.700~0.753)
Lyman/log-logistic	0.613 (0.583~0.643)	0.569 (0.545~0.594)

Independent test on 25 newly treated patients

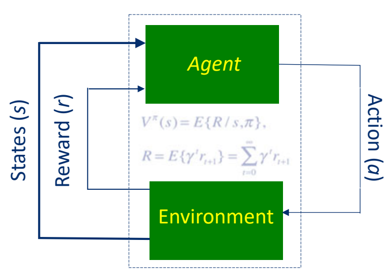
C-index (95%CI)	RP2	LC
NN-composite	0.692	0.721
NN-DVH	0.684	0.706
Lyman/log-logistic	0.588	0.573

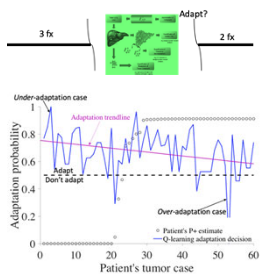


Cui et al, IJROBP, 2021

13

How to optimize RT adaptation decision with ML?

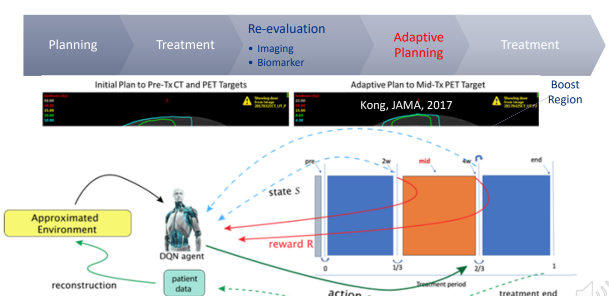




El Naqa, ASTRO (Best of physics), 2016

14

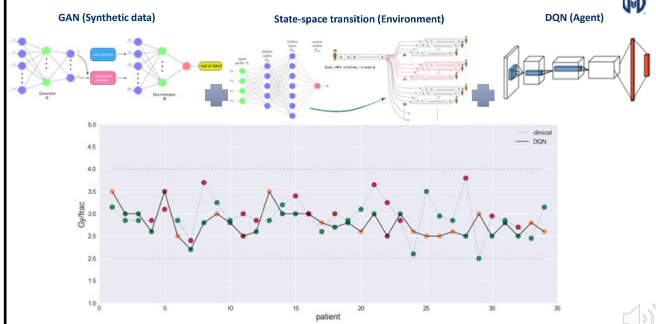
Adaptive Radiation Oncology Decision Making with Deep Learning I



Tseng, Med Phys, 2017 (Farrington Daniels Award in MP)

15

Adaptive Radiation Oncology Decision Making with Deep Learning II



Tseng, Med Phys, 2017 (Farrington Daniels Award)

16

Quality assurance for AI/ML application in the clinic

cceptance Testing
- To ensure that the ML tool meets all applicable safety and performance standards (prediction) and that it meets contractual specifications
- Manufacturer includes an acceptance test procedure with the ML tool
  - Selection of evaluation endpoint and definition of performance criteria (e.g., AUC);
  - Selection of a benchmark data

ommissioning
- The process whereby the needed tool-specific data/parameters are acquired and operational procedures are defined
- May include:
  - Training data collection
  - Developing procedures
  - User training before first use

uality Assurance (QA)
- Effort to ensure treatments are given accurately, safely and efficiently according to established tests and evaluations

ontinuing Quality Improvement (CQI)
- Effort that seeks to make treatments and operations better by recognizing current weaknesses in the program, anticipating problems before they happen, streamlining tasks and responding to changes in practice

Type of ML Application	QA Considerations for the Current State			
	Acceptance Testing	Commissioning	Routine QA	Risk being mitigated
ML replaces human tasks: Linear Accelerator QA	Confirm functionality with sample QA data (Ritter *et al.*, 2018)	Evaluate ML against current clinic standards (Klein *et al.*, 2009) Test limits of analytics such as by inserting errors into delivery tests and/or datasets for analysis, e.g., intentional leaf offset present in the measurement result but missing in the delivery file Document situations where the software passes and fails Document situations where results differ by >5%	Frequency: monthly Monitor software settings for analysis Repeat analysis of a subset of the commissioning dataset (e.g., dynamic leaf gap) including one at the limit Expect identical results unless the software has changed. If software has changed, determine if a new baseline is needed Evaluate against a subset of the manual analysis for software update Review trends	Confirm that the analysis is performed correctly to avoid the hazards of expectation bias
ML supplemental to human tasks: Treatment planning	Confirm functionality with vendor supplied treatment plans Define scope of ML for planning	Evaluate behavior against appropriate portions of original TPS commissioning results (if available) (Fraass *et al.*, 1998) Are clinical goals met? Is the agreement within +/- 5% for key metrics such as mean dose for targets and max dose to a volume (e.g. 1cc)? Evaluate ML tools for a range of body sites and have site-specific rollout of techniques for at least a limited number of body sites Evaluate permissions of different user types for applying ML techniques (e.g., physicist vs dosimetrist) Have different users perform the same test case – results within 5%? Establish procedures for quality control steps post-application of ML, e.g., MD and physicist review of final dose distribution	Repeat analysis of a subset of the commissioning dataset (e.g., dynamic leaf gap) including one at the limit Monitor key dosimetric results from ML techniques using Big Data Analytical tools where available by body site: e.g. target coverage and maximum dose to a volume (e.g. 1cc) for OARs (Mayo *et al.*, 2017) Add extra scrutiny on key metrics for the first 5 patients per body site	Monitor for any un-intentional shift in clinical practice due to settings in the ML algorithm Maintain evaluation of plan against MD provided goals (planning objectives)(Evans *et al.*, 2016; Marks *et al.*, 2013)
ML/AI enhances human tasks: Patient workflow such as preparation for optimization	Confirm functionality and understand the scope of what is automated	Define if ML tools will be applied and implemented for all patients or by body site Create a commissioning dataset which includes manual preparation of the plan for optimization and automated preparation Confirm reasonably concordant results between human and automated creation Inspect the overlay of human vs automated volumes to confirm expansions are correct Verify volumes for optimization are within 5% or 2 cc (for optic and other small structures)	Repeat a subset of the commissioning dataset Confirm derivative structures such as optimization structures are consistent with those by humans (monthly) Confirm that quality control steps post-application remain in place such as review of the final dose distribution by MD and physicist	Risk being mitigated is an incorrect expansion from target or OAR volumes to create optimization structures for dose coverage or sparing respectively Maintain evaluation of plan against MD provided goals (planning directive)(Evans *et al.*, 2016; Marks *et al.*, 2013)
ML Additive: Decision making (El Naqa *et al.*, 2018a)	Evaluate with vendor-supplied dataset Define size of training and testing datasets	Partner with physicians to determine which disease types and staging are appropriate for the algorithm Assess baseline variation in clinical practice among physicians within a practice, within a registry, or via publications before implementation Assess sensitivity of the output of algorithms with training sets across the spectrum of limited variability to significant variability Is the algorithm supporting implementation of a national practice standard? Is the algorithm being used to apply new science in a clinical trial?	Confirm that the input and expected output are consistent with the intent of the practice Assess the frequency of patient-type to determine how often the training dataset should be updated Monitor the relationship between decisions with prior practice using Big Data Analytical tools where available by body site	

El Naqa, Moran, Ten Haken, The Modern Technology of Radiation Oncology, V4, Van Dyke

17

ML Accuracy versus interpretability

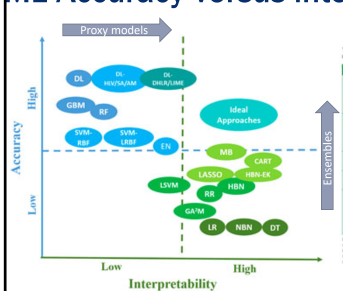


Table 1. The evaluation of the accuracy (A), interpretability (I) and explainability (E) of ML approaches in radiation outcomes prediction

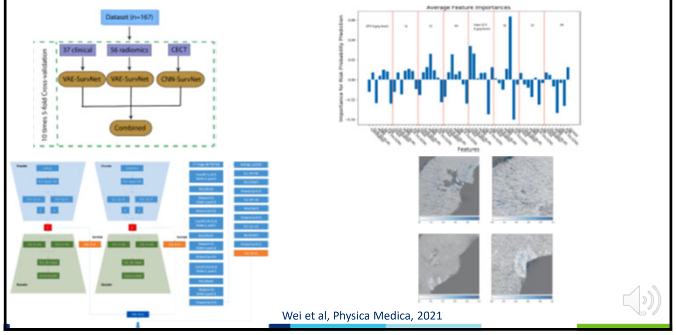
Basic ML	Type	A	I	E	Improved ML	Type	A	I	E
Logistic regression[20,21]	IP	*	****	***	GA²M[49]	IP	**	***	**
					Ridge Regression[22]	IP	**	**	*
					LASSO[23]	IP	**	***	**
					Elastic Net[8,24]	IP	***	**	*
Decision tree [24,30,31]	IP	**	*****	*****	CART[32]	IP	***	****	*****
					Random Forests[7]	NIP	****	*	NA
					GBM[8,33]	NIP	****	*	NA
					MediBoost[8,34]	IP	****	**	*
Naive BN [35,37]	IP	*	****	****	HBN[38,40]	IP	**	***	**
					HBN-EK[41]	IP	**	****	***
Linear SVM [24]	NIP	**	**	*	SVM-RBF[43]	NIP	***	*	NA
					SVM-LRBF[44]	NIP	***	**	*
Deep learning[49,50]	NIP	****	*	NA	DL-HLV [48,55,56]	NIP	*****	**	NA
					DL-SA[52,57] /AM[58,60]	NIP	*****	**	NA
					DL-DHLR[61–63]	NIP	*****	***	NA
					DL-LIME[69]	NIP	*****	***	NA

BN, Bayesian network; CART, classification and regression tree; DHLR, disentangled hidden layer representation; DL-AM, deep learning withattention mechanisms; DL-HLV, deep learning withcombination of handcrafted features and latent variables; GBM, gradient boosting machine; HBN, hierarchical Bayesian network; HBN-EK, hierarchical Bayesiannetwork with expert knowledge; HLV, handcrafted features and latent variables; IP, interpretable; LASSO, least absolute shrinkage and selection operator; LIME, local interpretable model-agnostic explanation; ML, machine learning; NIP, non-interpretable; SVM, support vector machine.

Luo, BJR-O, 2019

18

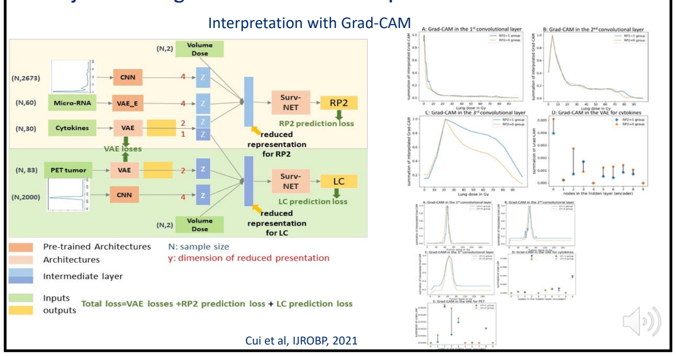
Deep Survival Radiomics model for Liver Cancer II



Wei et al, Physica Medica, 2021

19

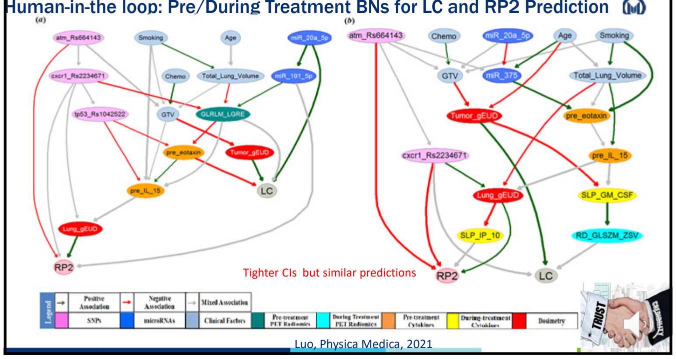
Multi-objective radiogenomics model with deep survival neural networks II



Cui et al, IJROBP, 2021

20

Human-in-the loop: Pre/During Treatment BNs for LC and RP2 Prediction



Luo, Physica Medica, 2021

21

Take home Messages

- Artificial intelligence/machine learning offers new opportunities to develop better understanding of oncology processes and improve its workflow and especially decision support systems
- Varying machine learning algorithms can be deployed. Deep learning methods can incorporate data representation and task learning in the same framework
- Collaboration between stakeholders (data scientists, clinicians, & biologists) will allow for safe and beneficial application of AI in biomedicine

22

THANK YOU!

23