Radiomics and Radiogenomics Modeling with Machine Learning

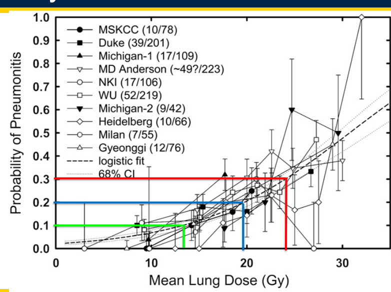
Issam El Naqa, PhD, DABR, FAAPM

Department of Machine Learning
Division of Quantitative Science
Moffitt Cancer Center
July 16th, 2020

1

Why 'omics outcomes modeling?

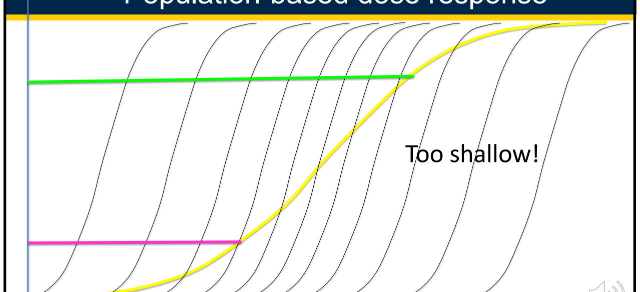
QUANTEC Symptomatic pneumonitis vs. mean lung dose



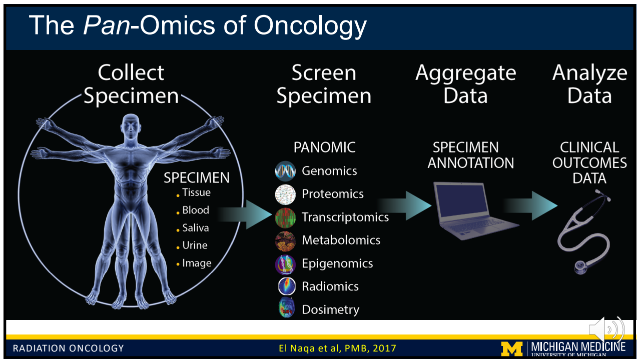
ΔP/ΔMLD < 2 %/Gy

2

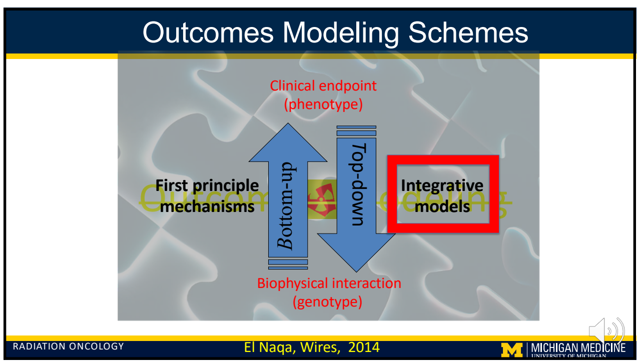
Population-based dose response



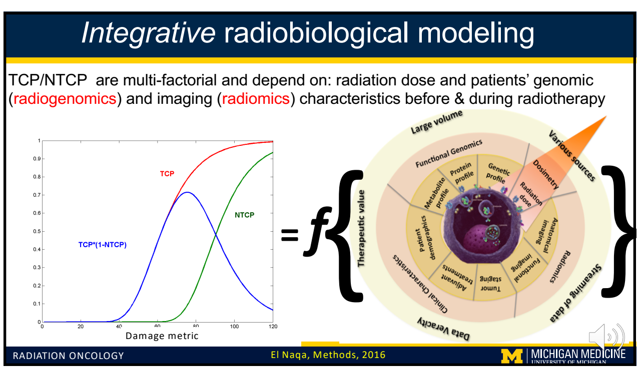
3


The Pan-Omics of Oncology
Collect Specimen
Screen Specimen
Aggregate Data
Analyze Data
SPECIMEN
Tissue
Blood
Saliva
Urine
Image
PANOMIC
Genomics
Proteomics
Transcriptomics
Metabolomics
Epigenomics
Radiomics
Dosimetry
SPECIMEN ANNOTATION
CLINICAL OUTCOMES DATA
RADIATION ONCOLOGY
El Naqa et al, PMB, 2017
MICHIGAN MEDICINE

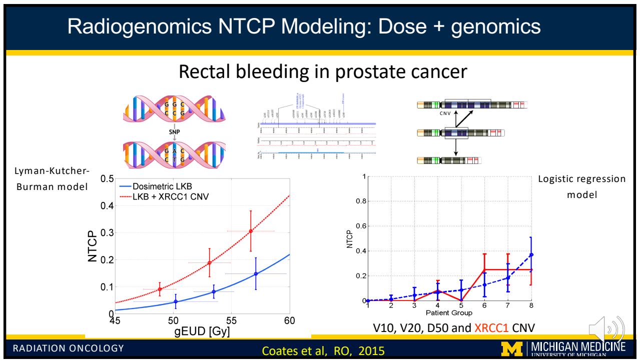
4


Outcomes Modeling Schemes
Clinical endpoint (phenotype)
First principle mechanisms
Bottom-up
Top-down
Integrative models
Biophysical interaction (genotype)
RADIATION ONCOLOGY
El Naqa, Wires, 2014
MICHIGAN MEDICINE

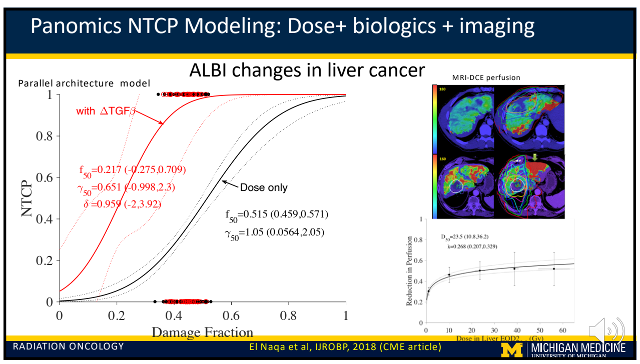
5


Integrative radiobiological modeling
TCP/NTCP are multi-factorial and depend on: radiation dose and patients' genomic (radiogenomics) and imaging (radiomics) characteristics before & during radiotherapy
TCP
NTCP
TCP*(1-NTCP)
Damage metric
= f{ }
Large volume
Various sources
Streaming of data
Data Veracity
Therapeutic value
Functional Genomics
Radiomics
Dosimetry
RADIATION ONCOLOGY
El Naqa, Methods, 2016
MICHIGAN MEDICINE

6



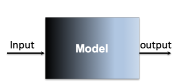
7



8

Outcome modeling by Machine learning (ML)

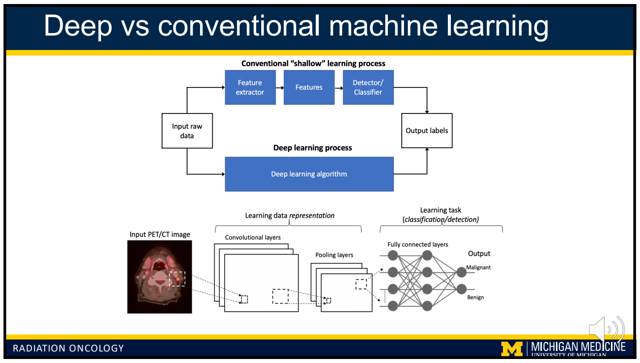
– *Generative* models
- Model class-conditional PDFs and prior probabilities (Bayesian networks, Markov models)
  – *To predict you need to know the system*



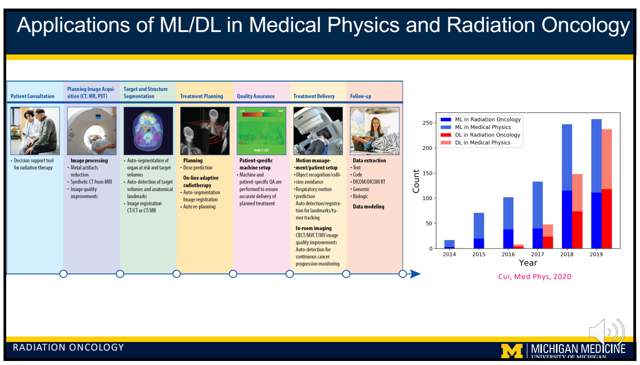
– *Discriminant* models
- Directly estimate posterior probabilities (logistic regression, neural networks, CNN, random forests, SVM)
  – *Predict without knowing the system*

RADIATION ONCOLOGY | MICHIGAN MEDICINE

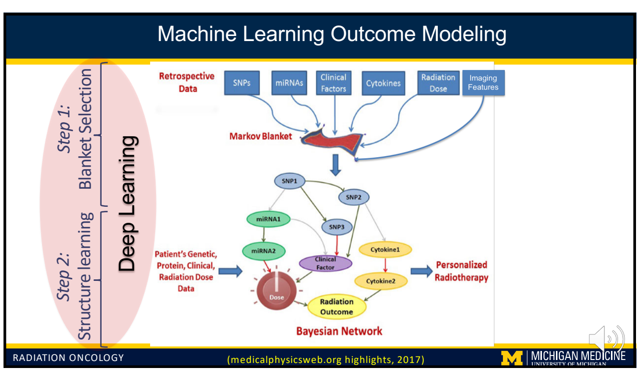
9


Deep vs conventional machine learning
Conventional "shallow" learning process
Feature extractor
Features
Detector/ Classifier
Input raw data
Output labels
Deep learning process
Deep learning algorithm
Learning data representation
Learning task (classification/detection)
Input PET/CT image
Convolutional layers
Pooling layers
Fully connected layers
Output
Malignant
Benign
RADIATION ONCOLOGY
MICHIGAN MEDICINE

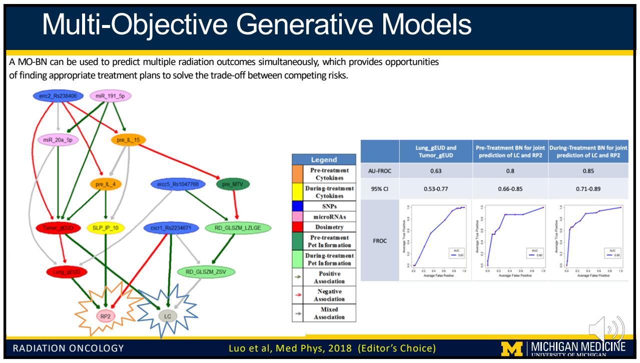
10


Applications of ML/DL in Medical Physics and Radiation Oncology
Patient Consultation
Planning Image Acquisition (CT, MR, PET)
Target and Structure Segmentation
Treatment Planning
Quality Assurance
Treatment Delivery
Follow-up
ML in Radiation Oncology
ML in Medical Physics
DL in Radiation Oncology
DL in Medical Physics
Count
Year
2014
2015
2016
2017
2018
2019
Cui, Med Phys, 2020
RADIATION ONCOLOGY
MICHIGAN MEDICINE

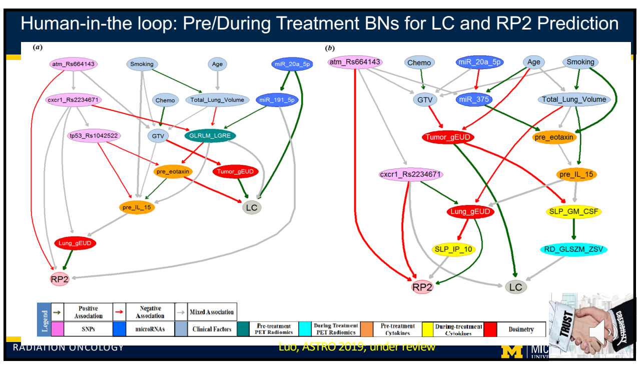
11


Machine Learning Outcome Modeling
Step 1: Blanket Selection
Step 2: Structure learning
Deep Learning
Retrospective Data
SNPs
miRNAs
Clinical Factors
Cytokines
Radiation Dose
Imaging Features
Markov Blanket
SNP1
SNP2
SNP3
miRNA1
miRNA2
Clinical Factor
Cytokine1
Cytokine2
Dose
Radiation Outcome
Patient's Genetic, Protein, Clinical, Radiation Dose Data
Personalized Radiotherapy
Bayesian Network
RADIATION ONCOLOGY
(medicalphysicsweb.org highlights, 2017)
MICHIGAN MEDICINE

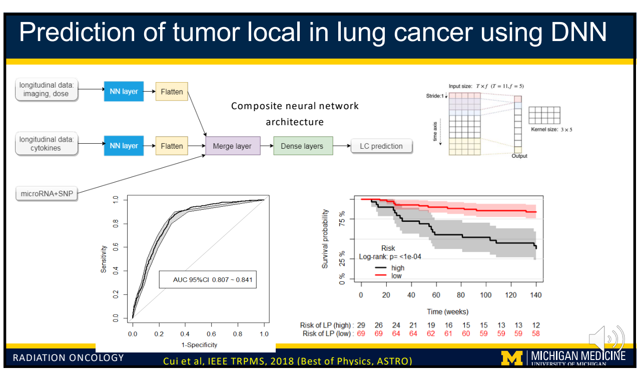
12


Multi-Objective Generative 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.
Legend
Pre-treatment Cytokines
During-treatment Cytokines
SNPs
microRNAs
Dosimetry
Pre-treatment Pet Information
During-treatment Pet Information
Positive Association
Negative Association
Mixed Association
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
RADIATION ONCOLOGY
Luo et al, Med Phys, 2018 (Editor's Choice)
MICHIGAN MEDICINE

13

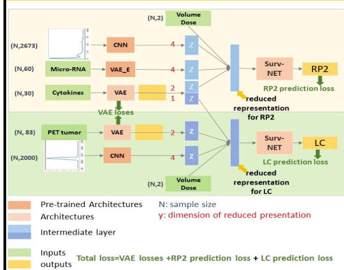

Human-in-the loop: Pre/During Treatment BNs for LC and RP2 Prediction
(a)
(b)
atm_Rs664143
Smoking
Age
miR_20a_5p
Chemo
GTV
miR_375
Total_Lung_Volume
Tumor_gEUD
pre_eotaxin
cxcr1_Rs2234671
pre_IL_15
Lung_gEUD
SLP_GM_CSF
SLP_IP_10
RD_GLSZM_ZSV
RP2
LC
Legend
Positive Association
Negative Association
Mixed Association
SNPs
microRNAs
Clinical Factors
Pre-treatment PET Radiomics
During Treatment PET Radiomics
Pre-treatment Cytokines
During-treatment Cytokines
Dosimetry
RADIATION ONCOLOGY
Luo, ASTRO 2019, under review

14


Prediction of tumor local in lung cancer using DNN
longitudinal data: imaging, dose
NN layer
Flatten
longitudinal data: cytokines
Merge layer
Dense layers
LC prediction
microRNA+SNP
Composite neural network architecture
AUC 95%CI 0.807 ~ 0.841
Sensitivity
1-Specificity
Survival probability
Risk
Log-rank: p= <1e-04
high
low
Time (weeks)
Risk of LP (high) : 29 26 24 21 19 16 15 15 13 13 12
Risk of LP (low) : 69 69 64 64 62 61 60 59 59 59 58
RADIATION ONCOLOGY
Cui et al, IEEE TRPMS, 2018 (Best of Physics, ASTRO)
MICHIGAN MEDICINE

15

Deep learning architectures for joint actuarial prediction of LC and RP2

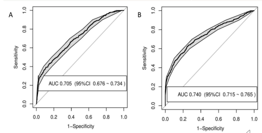


20 times of 5-fold cross validations

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



RADIATION ONCOLOGY Cui et al, AAPM 2020 MICHIGAN MEDICINE

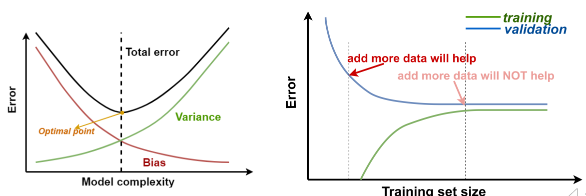
16

How to build an ML/DL model?

Introduction to Machine and Deep Learning for Medical Physicists

Sunan Cui, Huan-Hsin Tseng, Julia Pakela, Randall K. Ten Haken, and Issam El Naqa

Department of Radiation Oncology, University of Michigan, Ann Arbor, MI 48103, USA



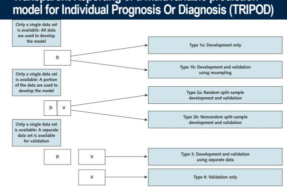
RADIATION ONCOLOGY Medical Physics Special Issue on Machine Learning MICHIGAN MEDICINE

17

ML validation

- Depending on the level of evidence
  - Selection appropriate learning algorithms
  - Validation and evaluation (TRIPOD criteria)
    - Internally (cross-validation schemes)
    - Externally (independent datasets)
  - Provide interpretation of machine learning prediction

Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD)



Analysis Type	Description
Type 1a	Development of a prediction model where predictive performance is then directly evaluated using exactly the same data (apparent performance).
Type 1b	Development of a prediction model using the entire data set, but then using resampling (e.g., bootstrapping or cross-validation) techniques to evaluate the performance and optimism of the developed model. Resampling techniques, generally referred to as "internal validation", are recommended as a prerequisite for prediction model development, particularly if data are limited (6, 14, 15).
Type 2a	The data are randomly split into 2 groups: one to develop the prediction model, and one to evaluate its predictive performance. This design is generally not recommended or better than type 1b, particularly in case of limited data, because it leads to lack of power during model developmen and validation (14, 15, 16).
Type 2b	The data are nonrandomly split (e.g., by location or time) into 2 groups: one to develop the prediction model and one to evaluate its predictive performance. Type 2b is a stronger design for evaluating model performance than type 2a, because allows for nonrandom variation between the 2 data sets (6, 13, 17).
Type 3	Development of a prediction model using 1 data set and an evaluation of its performance on separate data (e.g., from a different study).
Type 4	The evaluation of the predictive performance of an existing (published) prediction model on separate data (13).

Types 3 and 4 are commonly referred to as "external validation studies." Arguably type 2b is as well, although it may be considered an intermediary between internal and external validation.

RADIATION ONCOLOGY MICHIGAN MEDICINE UNIVERSITY OF MICHIGAN

18

ML Accuracy versus interpretability

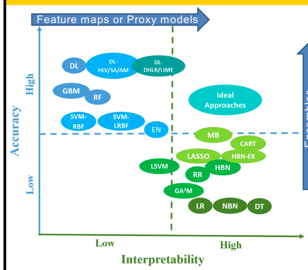


Table 1. The evaluation of the accuracy (**A**), interpretability (**I**) and explainability (**E**) of ML approaches in radiation outcomes prediction

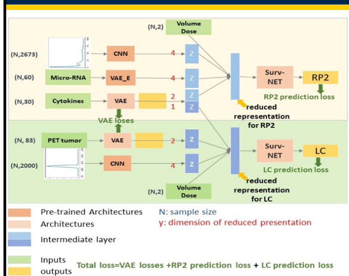
Basic ML	Type	A	I	E	Improved ML	Type	A	I	E
Logistic regression[20,21]	IP	*	****	***	GA²M[68]	IP	**	***	**
					Ridge Regression[22]	IP	**	**	*
					LASSO[23]	IP	**	***	**
					Elastic Net[9,24]	IP	***	**	*
Decision tree[24,30,31]	IP	**	*****	*****	CART[32]	IP	***	****	*****
					Random Forests[7]	NIP	****	*	NA
					GBM[9,33]	NIP	****	*	NA
					MediBoost[9,34]	IP	****	**	*
Naïve 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[59,60]	NIP	*****	**	NA
					DL-DHLR[61–63]	NIP	*****	***	NA
					DL-LIME[69]	NIP	*****	***	NA

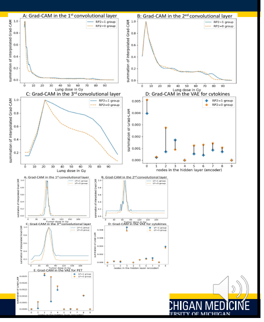
BN, Bayesian network; CART, classification and regression tree; DHLR, disentangled hidden layer representation; DL-AM, deep learning withattention mechanisms; DL-HLV, deep learning withcombination of handcrafted features and latent variables; GBM, gradient boosting machine; HBN, hierarchical Bayesian 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.

RADIATION ONCOLOGY | Luo, BJR-O, 2019 | MICHIGAN MEDICINE UNIVERSITY OF MICHIGAN

19

Interpretation of deep learning architectures for prediction of LC and RP2





RADIATION ONCOLOGY | Cui et al, AAPM 2020 | MICHIGAN MEDICINE

20

Conclusions

- Radiomics and radiogenomics offer new opportunities to develop better TCP/NTCP models and for personalizing radiotherapy
- Machine/deep learning techniques can improve feature selection and statistical learning in radiomics/radiogenomics analytics and modeling radiotherapy outcomes
- Main challenges for radiomics/radiogenomics modeling
  - Harmonization and optimization of data integration methods
  - Uncertainties in data and model building schemes
  - Proper validation (TRIPOD criteria) and robustness for clinical decision support
  - Better interpretation of radiomics/radiogenomics models is still lagging

RADIATION ONCOLOGY | MICHIGAN MEDICINE UNIVERSITY OF MICHIGAN

21

Machine Learning in Radiation Oncology
A GUIDE TO OUTCOME MODELING IN RADIOTHERAPY AND ONCOLOGY
MEDICAL PHYSICS
Emerging Developments and Practices in Oncology
WILEY
THANK YOU!
James Balter. PhD
Yue Cao, PhD
Paul Carson, PhD
Kyle Cuneo, MD
Joseph Deasy, PhD
Yuni Dewaraja, PhD
Ivo Dinov, PhD
Robert Gillies, PhD
Mike Green, MD, PhD
Judy Jin, PhD
Shruti Jolly, MD
Theodore Lawrence, MD PhD
Marc Kessler, PhD
Daniel L. McShan, PhD
Martha M. Matuszak, PhD
Chuck Mayo, PhD
Michelle Mierzwa, MD
Jean Moran, PhD
Eduardo Moros, PhD
Morand Piert, MD
Dipankar Ray, PhD
Al Rehemtulla, PhD
Ben Rosen, PhD
Jan Seuntjens, PhD
Randall K. Ten Haken, PhD
Xueding Wang, PhD
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
National Institutes of Health
RADIATION ONCOLOGY

22