



# Big Data and AI for Compliance Prediction

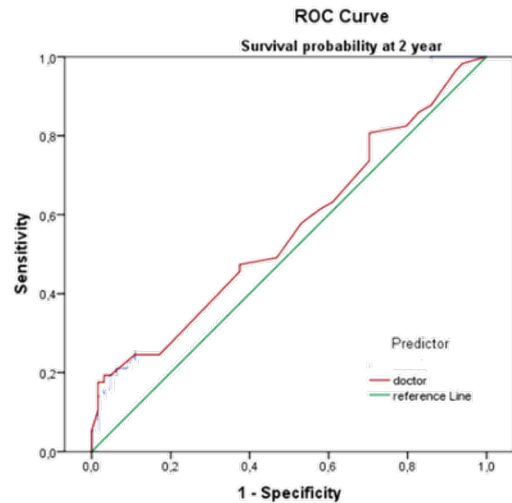
---

Andre Dekker | Medical Physicist | Professor of Clinical Data Science  
Maastric Clinic, Maastricht University, Maastricht UMC+

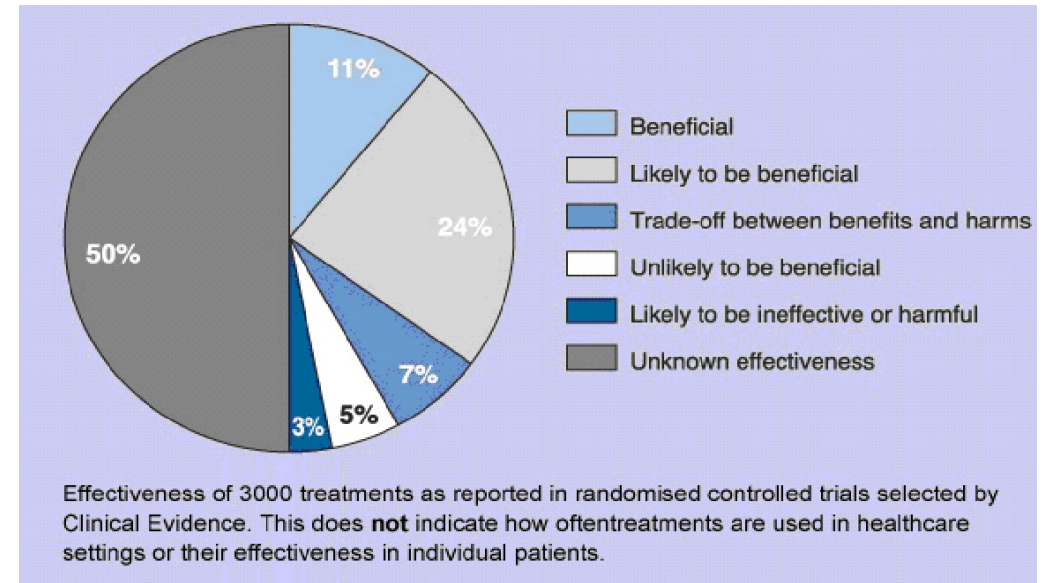
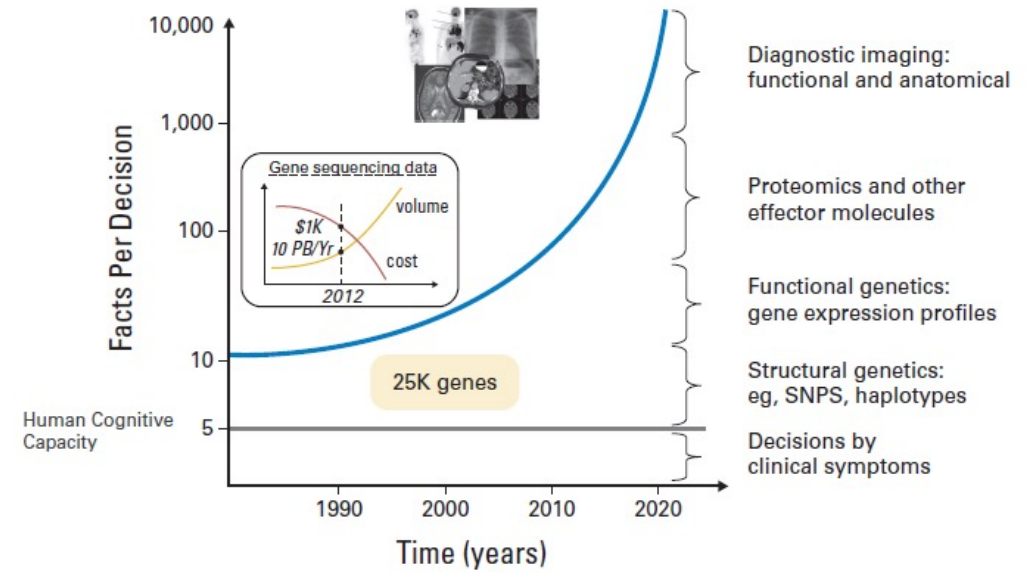
# Disclosures

- Research collaborations incl. funding, consultancy and speaker honoraria
  - Pharma: Roche, Johnson & Johnson, Bristol-Myers Squibb
  - MedTech: Varian Medical Systems, Siemens, Philips, Sohard, Mirada Medical, ptTheragnostics, OncoRadiomics
  - Health insurance: CZ Health Insurance
- Spin-offs and commercial ventures
  - MAASTRO Innovations B.V.
  - Medical Data Works B.V.
- Various patents on medical machine learning & Radiomics
- Public research funding
  - Radiomics (USA-NIH/U01CA143062),
  - duCAT&Strategy (NL-STW)
  - CloudAtlas, DART&Decide, SeDI (EU-EUROSTARS)
  - BIONIC, TRAIN ELIXIR (NL-NWO)
  - PROTRAIT&TraIT2HealthRI (NL-KWF)
  - Data4LifeSciences (NL-NFU)
  - Digital Society Agenda – Health&Well-Being (NL-VSNU)

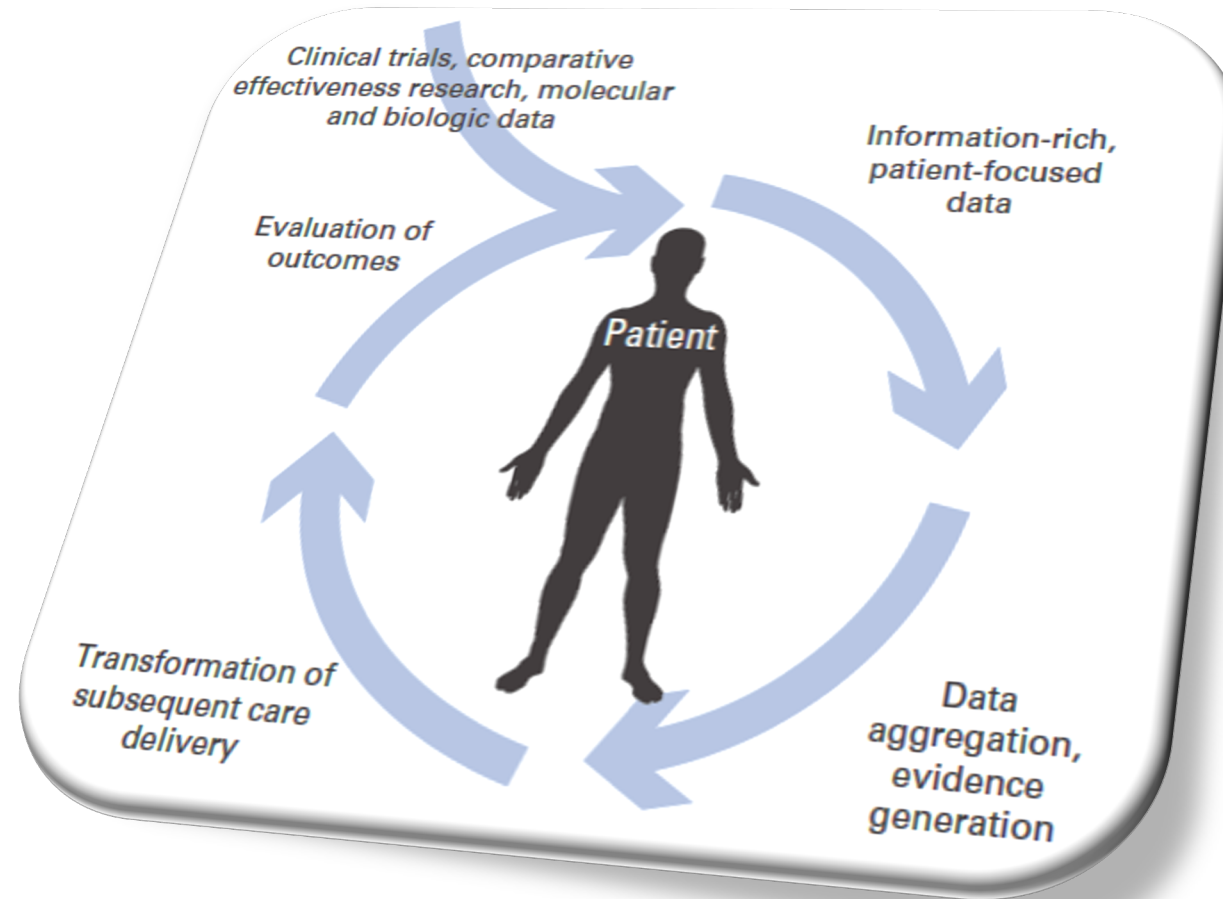
# General rationale Big Data & AI Prediction of Outcomes



NSCLC (Lung Cancer)  
 2 year survival  
 158 patients  
 5 MDs  
 Prospective  
**AUC: 0.56 == flipping a coin**



# General rationale Big Data & AI – Learning Health Care



# Elderly patient

- Elderly not often included in trials, not well known how they respond to RT
- Elderly patients more often non-compliant
  - In trials (CONVERT): 34% vs 13%
  - In clinic
- Many factors that may cause non-compliance in elderly (multimorbidity, frailty, weaker immune system, social isolation)
- Age is often in factor in our models predicting toxicity
  - “Baseline” xerostomia & dysphagia (head & neck cancer)
  - Radiation pneumonitis (lung cancer)
  - Cardiac toxicity (breast cancer)

Table 2. Radiotherapy Compliance per Age Group

Arm (N)	Dose (Gy)			No. Fractions			
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
BD (249)	<44	44-46	>46	<28	28-29	30	>30
<70 (220)	1 (0.4)	216 (98)	3 (1)	10 (5)	18 (8)	191 (87)	1 (0.4)
≥70 (29)	0 (0)	29 (100)	0 (0)	2 (7)	5 (17)	22 (76)	0 (0)
OD (240)	<60	60-62	64-68	<30	30-32	33	>33
<70 (202)	17 (8)	16 (8)	169 (84)	13 (6)	23 (11)	165 (82)	1 (0.5)
≥70 (38)	5 (13)	3 (8)	30 (79)	3 (8)	8 (21)	27 (71)	0 (0)

13%  
24%

BD, twice-daily; OD, once-daily.

Factors	Noncompliant patients, n (%)	Compliant patients, n (%)	P
Age (years)			
20-40	10% 4 (15.4)	55 (31.1)	0.07
40-60	12 (46.1)	86 (48.6)	
>60	22% 10 (38.5)	36 (20.3)	
Gender			
Male	22 (84.6)	94 (53.1)	0.002
Female	4 (15.4)	83 (46.9)	
Primary tumor site			
Head and neck	22 (84.6)	116 (65.5)	0.14
Cervix	3 (11.5)	39 (22.1)	
Breast	1 (3.9)	22 (12.4)	
AJCC stage			
I	0	4 (2.3)	0.004
II	2 (7.7)	73 (41.2)	
III	20 (76.9)	75 (42.4)	
IV	4 (15.4)	25 (14.1)	
CCRT			
Yes	25 (96.1)	93 (52.5)	<0.001
No	1 (3.9)	84 (47.5)	
Distance (km)			
<50	7 (26.9)	94 (53.1)	0.03
50-100	10 (38.5)	53 (30)	
>100	9 (34.6)	30 (16.9)	
Finance			
Paid	17 (65.4)	123 (69.5)	0.67
Free	9 (34.6)	54 (30.5)	

AJCC=American Joint Committee on Cancer; CCRT=Concurrent chemo-radiotherapy

# Elderly patient

- Age is often in factor in our models predicting toxicity
  - “Baseline” xerostomia & dysphagia (head & neck cancer)
  - Radiation pneumonitis, dysphagia (lung cancer)
  - Cardiac toxicity (breast cancer)

## Dyspnea Model Input

Age (36-90 years):

60

Nicotine use:

No/ex  Yes

WHO-PS:

1

FEV<sub>1</sub> (20-140%):

80

MLD (2-23Gy):

15

## Output Dyspnea Model

Probability to develop acute severe ( $\geq$  grade 2) dyspnea: 18%

95% Confidence interval: 12% - 25%

**Interpretation:** If there would be a group of 100 patients with the same characteristics as this individual patient, 18 patients would develop severe dyspnea ( $\geq$  grade 2) after the radiotherapy treatment. Due to the fact that a model can never be completely the same as the "real world", the number 18 could be lower or higher, but 18 is the most likely value. The 95% confidence interval indicates that the value will lie between 12 and 25 in 95% of the times that you would

## Dyspnea Model Input

Age (36-90 years):

80

Nicotine use:

No/ex  Yes

WHO-PS:

1

FEV<sub>1</sub> (20-140%):

80

MLD (2-23Gy):

15

## Output Dyspnea Model

Probability to develop acute severe ( $\geq$  grade 2) dyspnea: 27%

95% Confidence interval: 19% - 37%

**Interpretation:** If there would be a group of 100 patients with the same characteristics as this individual patient, 27 patients would develop severe dyspnea ( $\geq$  grade 2) after the radiotherapy treatment. Due to the fact that a model can never be completely the same as the "real world", the number 27 could be lower or higher, but 27 is the most likely value. The 95% confidence interval indicates that the value will lie between 19 and 37 in 95% of the times that you would

# First try in predicting compliance

- Doctors find it hard to predict toxicities
- Aim: A simple, transparent model (decision tree) that can predict compliance in elderly patients receiving RT
- Accepted for publication in *Frontiers in Oncology*

**Table 2**

Comparison of doctors' versus models' predictions.

	Outcome	Doctors'	
		AUC	95% CI
Timepoint 1	Dead within 2 years	0.56	0.46–0.67
	Dyspnea	0.59	0.44–0.74
	Dysphagia	0.52	0.39–0.66
Timepoint 2	Dead within 2 years	0.56	0.36–0.75
	Dyspnea	0.61	0.35–0.88
	Dysphagia	0.64	0.34–0.83

\* *p*-Value assessed with DeLong's test for two correlated ROC curves.

# Dataset

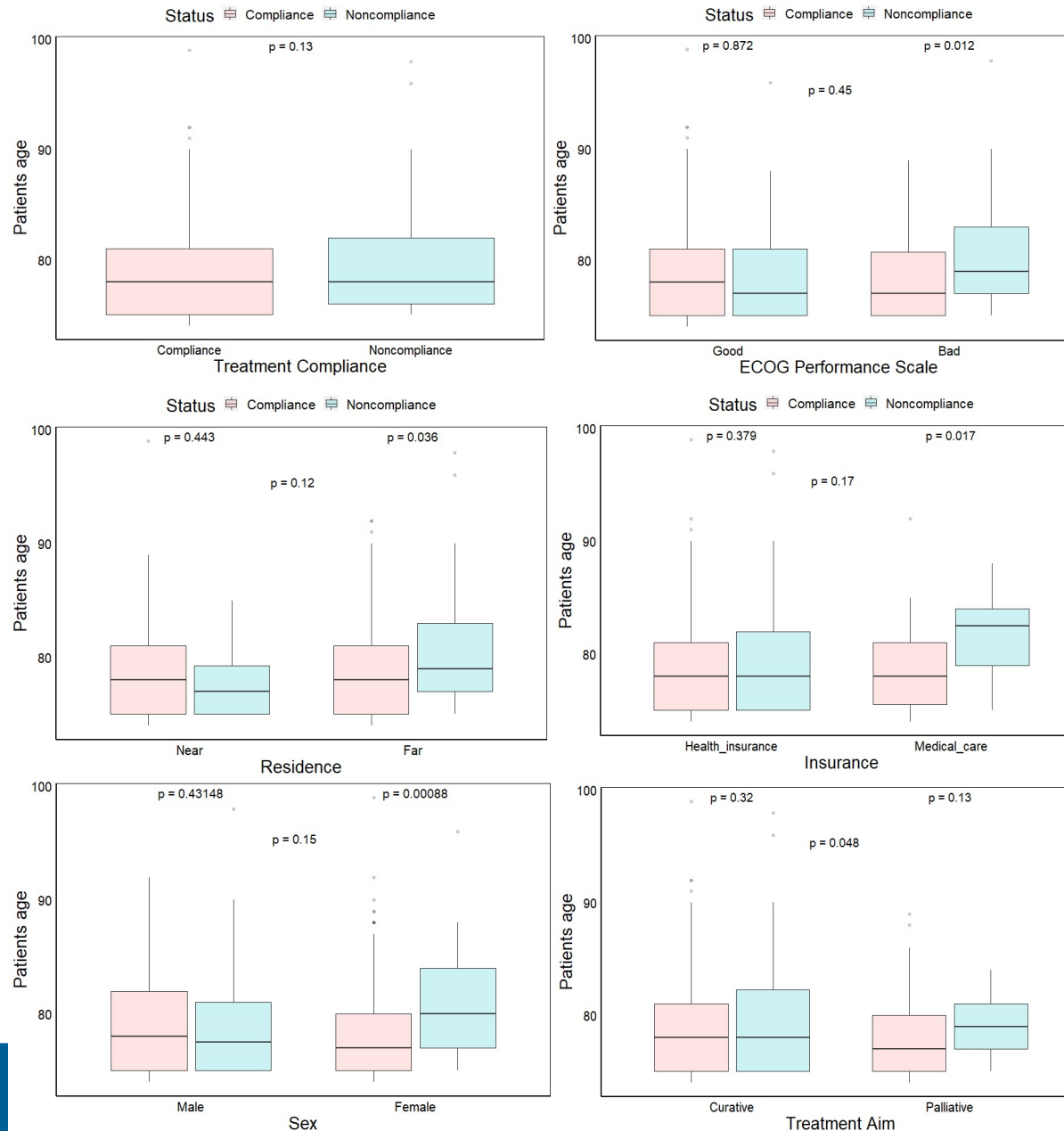
- Gil Medical Centre, Seoul, Korea
- 789 patients
- Median age 78, Range 74-99
- Radiotherapy
- Jan 2005 – Jan 2017
- Compliance == Completion of prescribed radiotherapy dose
- Noncompliance == Discontinuation of therapy against physician advice or consent
- Decision tree with Internal validation (Bootstrap, TRIPOD 2a)
- Considered predictive factors
  - Age
  - Gender
  - Eastern Cooperative Oncology Group (ECOG) Performance Status
  - Distance from home to radiotherapy center (residence)
  - Radiotherapy aim
  - Cancer type
  - Health insurance status (surrogate financial status)



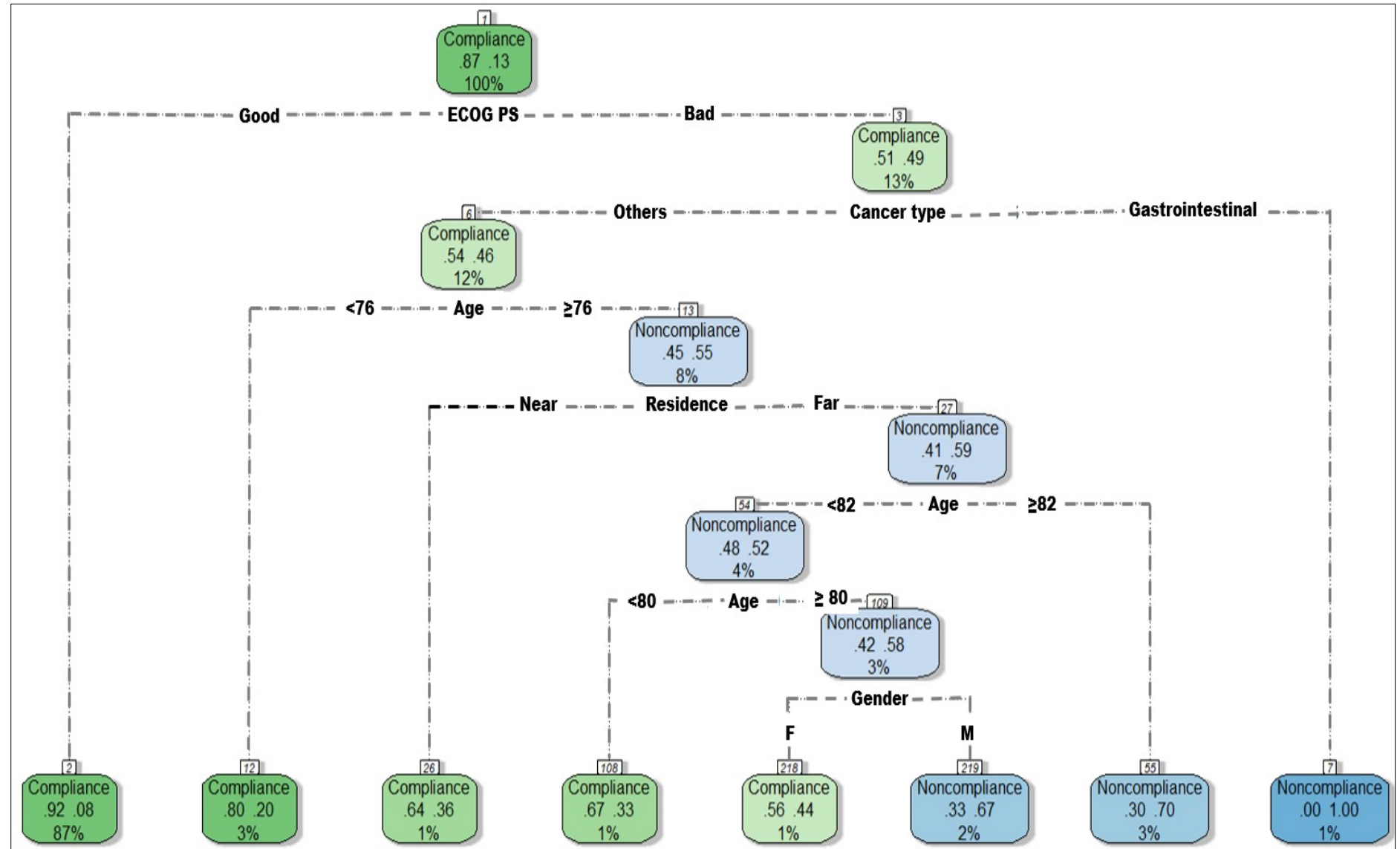
# Dataset

Variable	Levels	Compliance	Noncompliance	Total
Age	Mean (SD)	78 (4)	78 (5)	78 (4)
Sex	Male	367 (84%)	68 (16%)	435 (55%)
	Female	317 (90%)	37 (10%)	354 (45%)
ECOG PS	Poor (2+)	50 (51%)	<b>49 (49%)</b>	99 (13%)
	Good (0-1)	324 (92%)	56 (8%)	690 (87%)
Residence	Far	502 (86%)	81 (14%)	583 (74%)
	Near	182 (88%)	24 (12%)	206 (26%)
Radiotherapy aim	Curative	547 (85%)	84 (15%)	631 (80%)
	Palliative	137 (85%)	21 (15%)	158 (20%)
Health insurance status	Free medical care	79 (89%)	10 (11%)	89 (11%)
	Health insurance	605 (86%)	95 (14%)	700 (89%)
Cancer type	Skin	18 (82%)	4 (18%)	22 (3%)
	Lung	148 (85%)	27 (15%)	175 (22%)
	Brain	18 (90%)	2 (10%)	20 (3%)
	Breast	36 (97%)	1 (3%)	37 (5%)
	Sarcoma	06 (86%)	1 (14%)	7 (<1%)
	Metastatic	112 (84%)	22 (16%)	134 (17%)
	Hematologic	23 (96%)	1 (4%)	24 (3%)
	Hepatobiliary	38 (83%)	8 (17%)	46 (6%)
	Head & Neck	47 (78%)	13 (22%)	60 (8%)
	Genitourinary	57 (95%)	3 (5%)	60 (8%)
	Gynecological	104 (92%)	9 (8%)	113 (14%)
	Gastrointestinal	77 (85%)	14 (15%)	91 (11%)
	<b>Total</b>		<b>684 (87%)</b>	<b>105 (13%)</b>

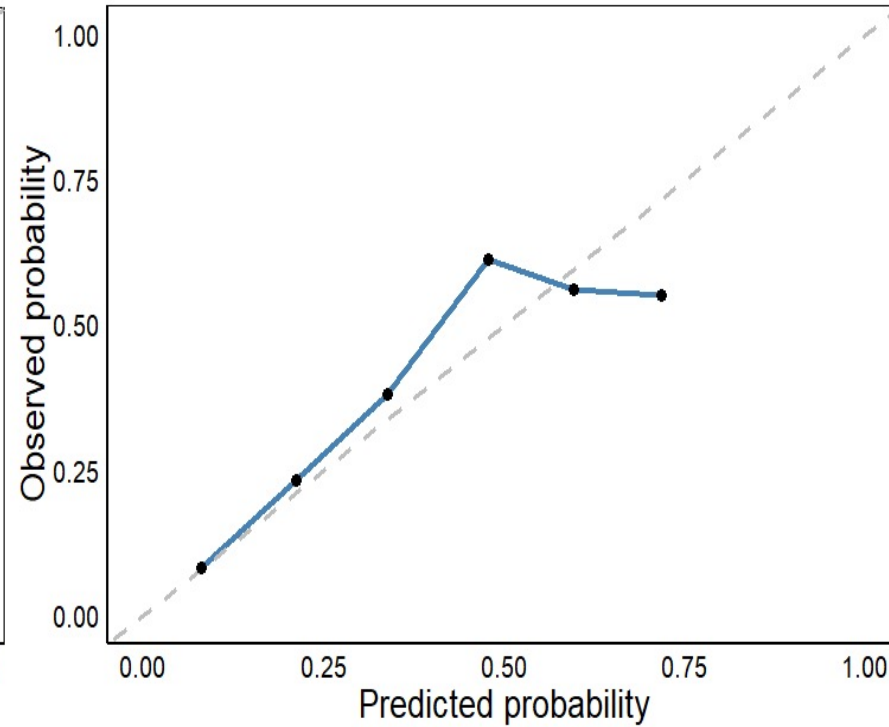
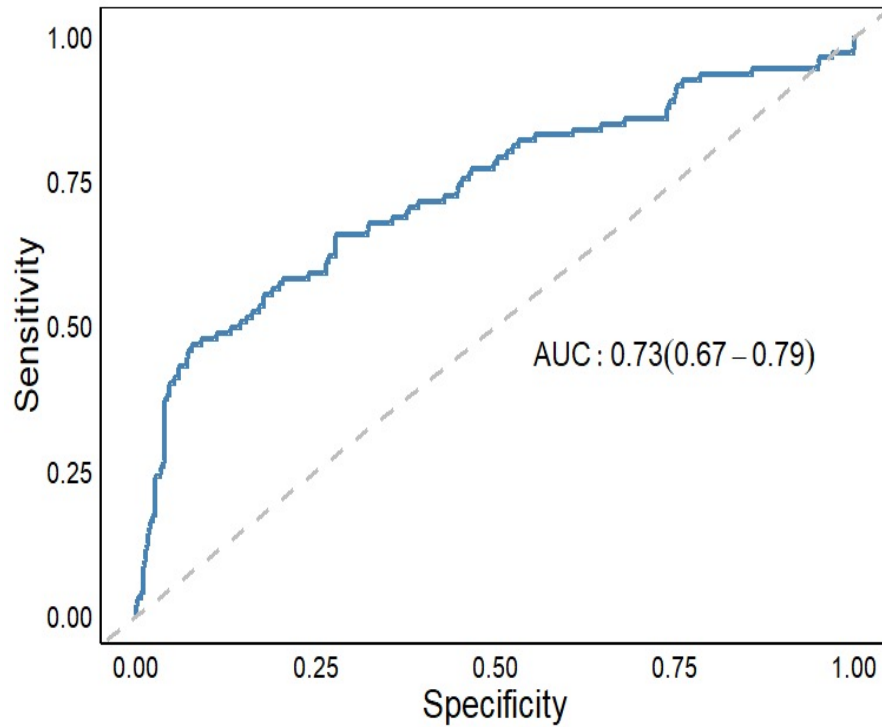
# Age



# Decision Tree



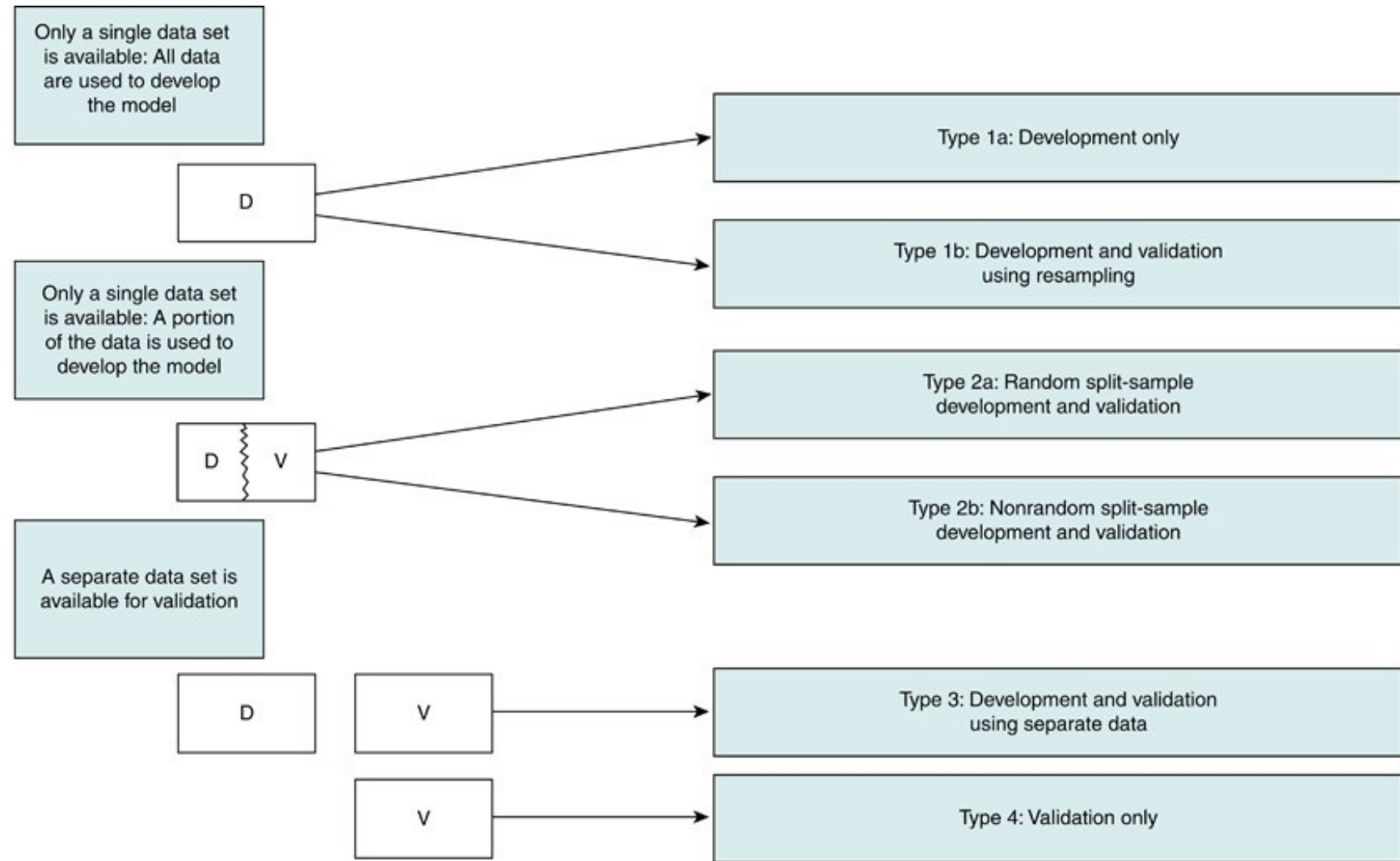
# Internal Bootstrap Validation



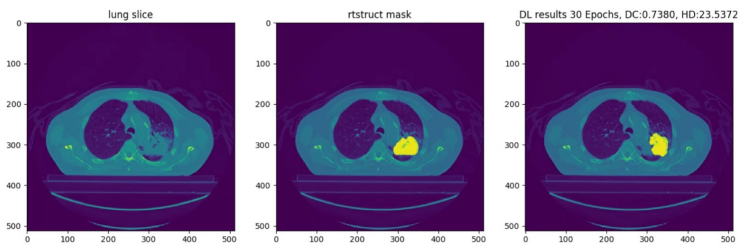
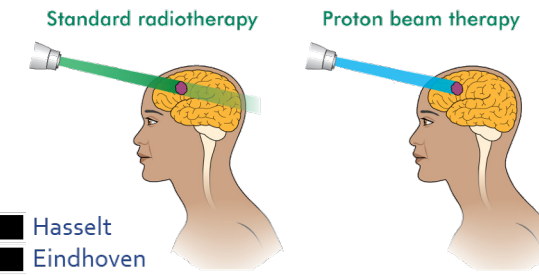
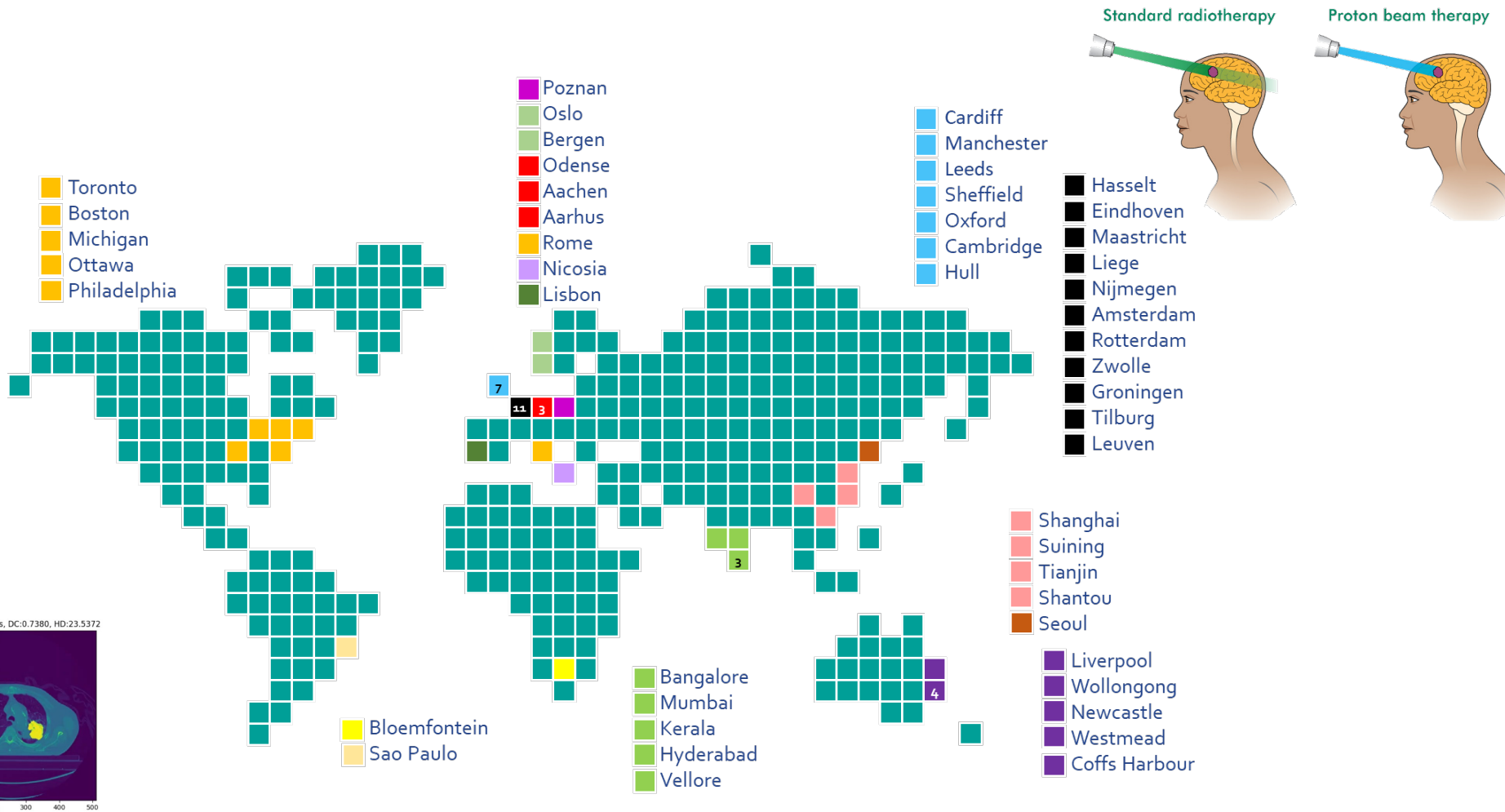
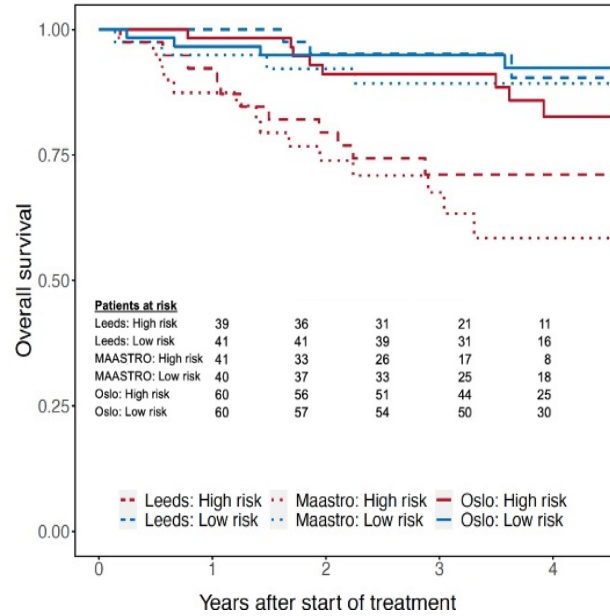
# Discussion / Limitations

- Internally validated study
- Predictive factors for compliance
  - Performance status
  - Cancer type
  - Distance to clinic
  - Age
- Possible actionable insights
  - SMS?
  - Hypofractionation?
  - Reconsider chemo or dose?

# Compliance of AI....



# CORAL: Community in Oncology for RApid Learning



#### Netherlands

- MAASTRO, Maastricht, Netherlands
- Radboudumc, Nijmegen, Netherlands
- Erasmus MC, Rotterdam, Netherlands
- Leiden UMC, Leiden, Netherlands
- Elizabeth Twee Steden Ziekenhuis, Tilburg, Netherlands
- Catharina Hospital, Eindhoven, Netherlands
- Isala Hospital, Zwolle, Netherlands
- NKI Amsterdam, Netherlands
- UMCG, Groningen, Netherlands
- IKNL, Utrecht, Netherlands

#### Europe

- Policlinico Gemelli & UCSC, Roma, Italy
- UH Ghent, Belgium
- UZ Leuven, Belgium
- Cardiff University & Velindre CC, Cardiff, UK
- CHU Liege, Belgium
- Uniklinikum Aachen, Germany
- LOC Genk/Hasselt, Belgium
- The Christie, Manchester, UK
- State Hospital, Rovigo, Italy
- St James Institute of Oncology, Leeds, UK
- U of Southern Denmark, Odense, Denmark
- Greater Poland Cancer Center, Poznan, Poland
- Oslo University Hospital, Oslo, Norway
- Aarhus Universitetshospital, Aarhus, Denmark
- Bank of Cyprus Oncology Center, Nicosia, Cyprus
- Weston Park Hospital, Sheffield, UK
- Hull University Teaching Hospitals NHS Trust, Hull, UK
- Addenbrookes' Hospital, Cambridge, UK
- Oxford University Hospitals NHS Foundation Trust, Oxford, UK
- Haukeland University Hospital, Bergen, Norway

#### Africa

- University of the Free State, Bloemfontein, South Africa

#### Asia

- Fudan Cancer Center, Shanghai, China
- CDAC, Pune, India
- Tata Memorial, Mumbai, India

- Suining Central Hospital, Suining, China
- HGC Oncology, Bangalore, India
- Kerala, Kerala, India
- Apollo Hospitals, Hyderabad, India
- CMC Vellore, Vellore, India
- MVRCC, Calicut, India
- Tianjin Medical University, Tianjin, China
- Cancer Hospital of Shantou University, Shantou, China

#### North America

- RTOG, Philadelphia, PA, USA
- MGH, BWH, Harvard, Boston, MA, USA
- University of Michigan, Ann Arbor, USA
- Princess Margaret CC, Canada
- Ottawa Hospital Research Institute, Ottawa, Canada

#### South America

- Albert Einstein, Sao Paulo, Brazil

#### Australia

- University of Sydney, Australia
- Westmead Hospital, Sydney, Australia
- Liverpool and Macarthur CC, Australia
- ICCA, Wollongong Australia
- Calvary Mater, Newcastle, Australia
- North Coast Cancer Institute, Coffs Harbour, Australia

#### Industry

- Varian, Palo Alto, CA, USA
- Philips, Bangalore, India
- Sohard GmbH, Fuerth, Germany
- Microsoft, Hyderabad, India
- Mirada Medical, Oxford, UK
- CZ Health Insurance, Tilburg, NL
- Siemens, Malvern, PA, USA
- Roche, Woerden, NL







# Thank you for your attention

---

Andre Dekker | Medical Physicist | Professor of Clinical Data Science  
Maastric Clinic, Maastricht University, Maastricht UMC+