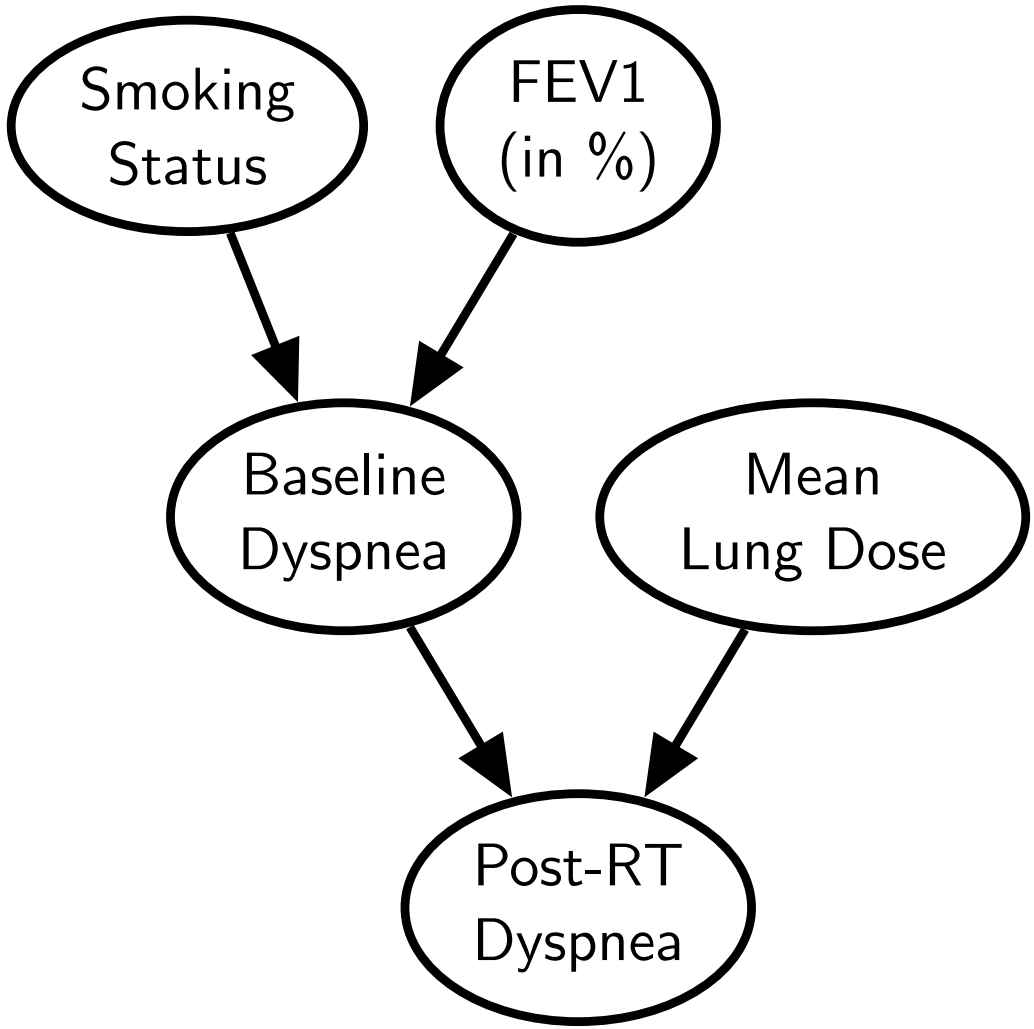
Expert knowledge and data-driven Bayesian Networks to predict post-RT dyspnea and 2-year survival

Timo M. Deist

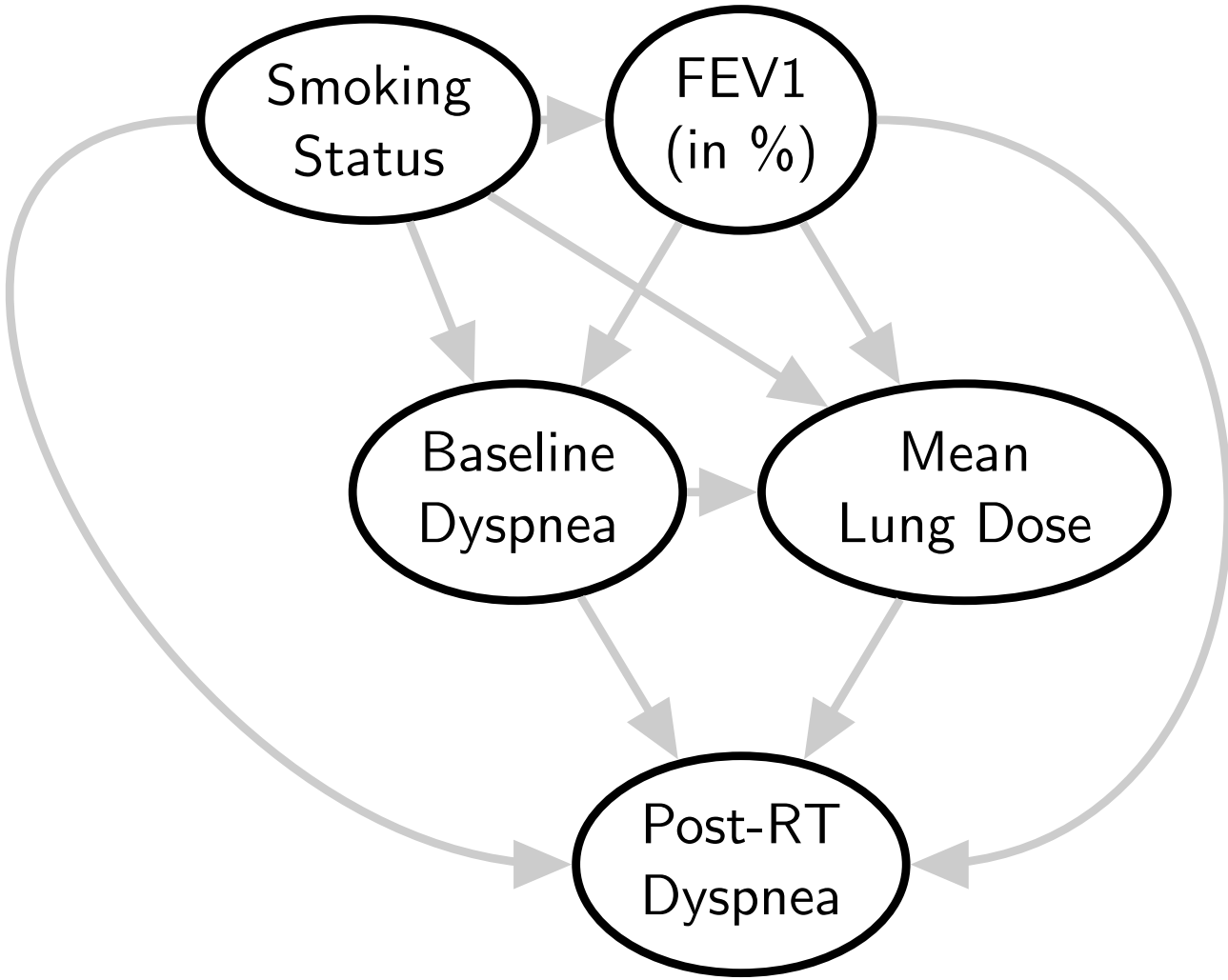
MAASTRO Clinic
Maastricht University
The Netherlands

16.02.2016

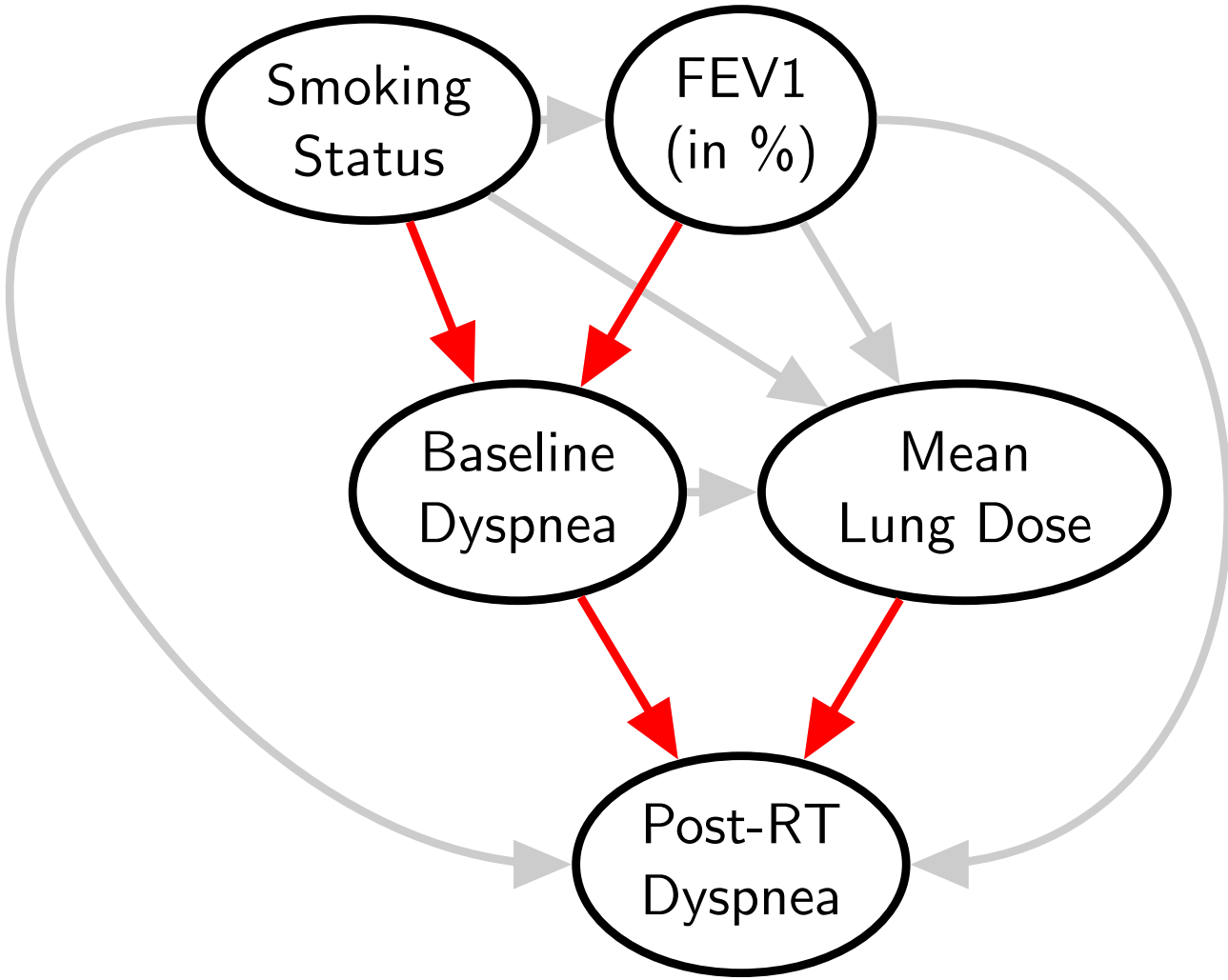
Example: Bayesian Network



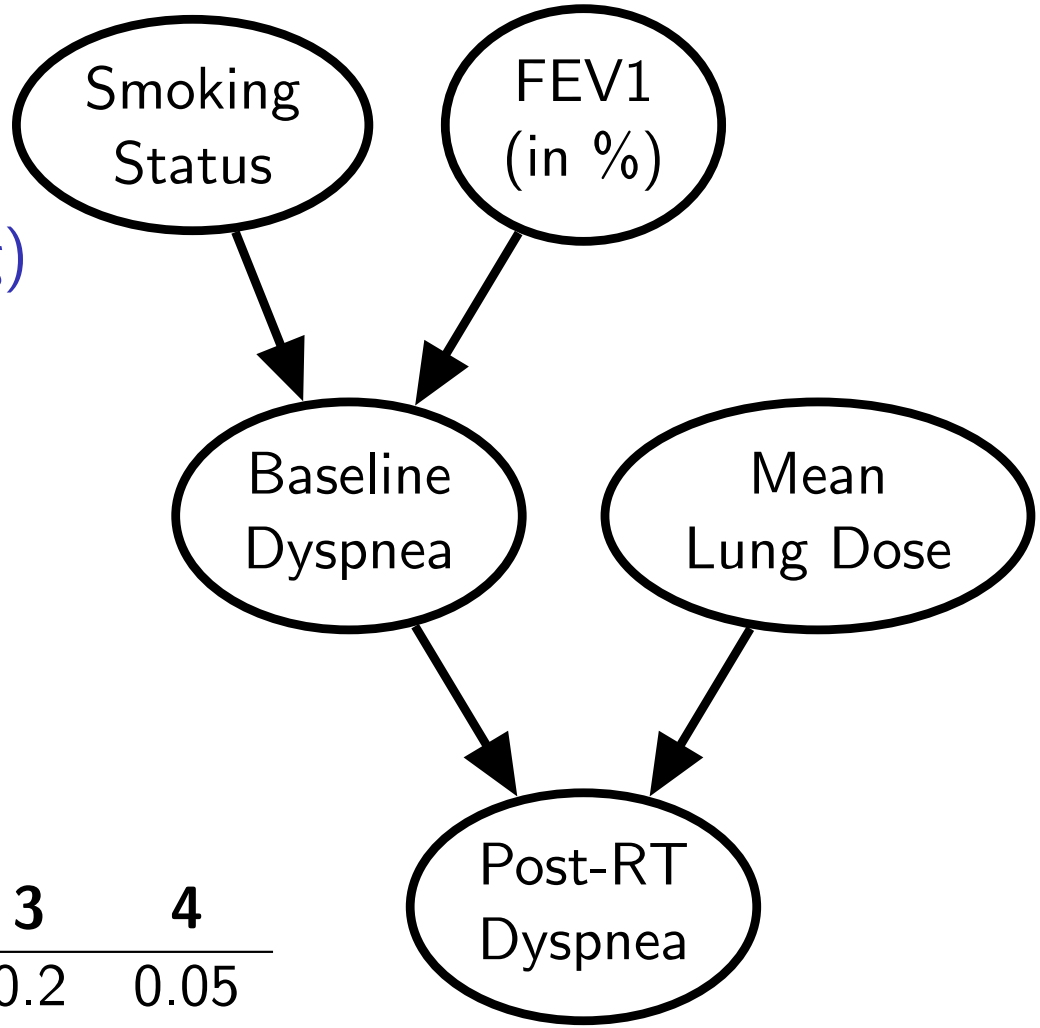
Step 1: Structure Learning



Step 1: Structure Learning



Step 2: Prediction Learning



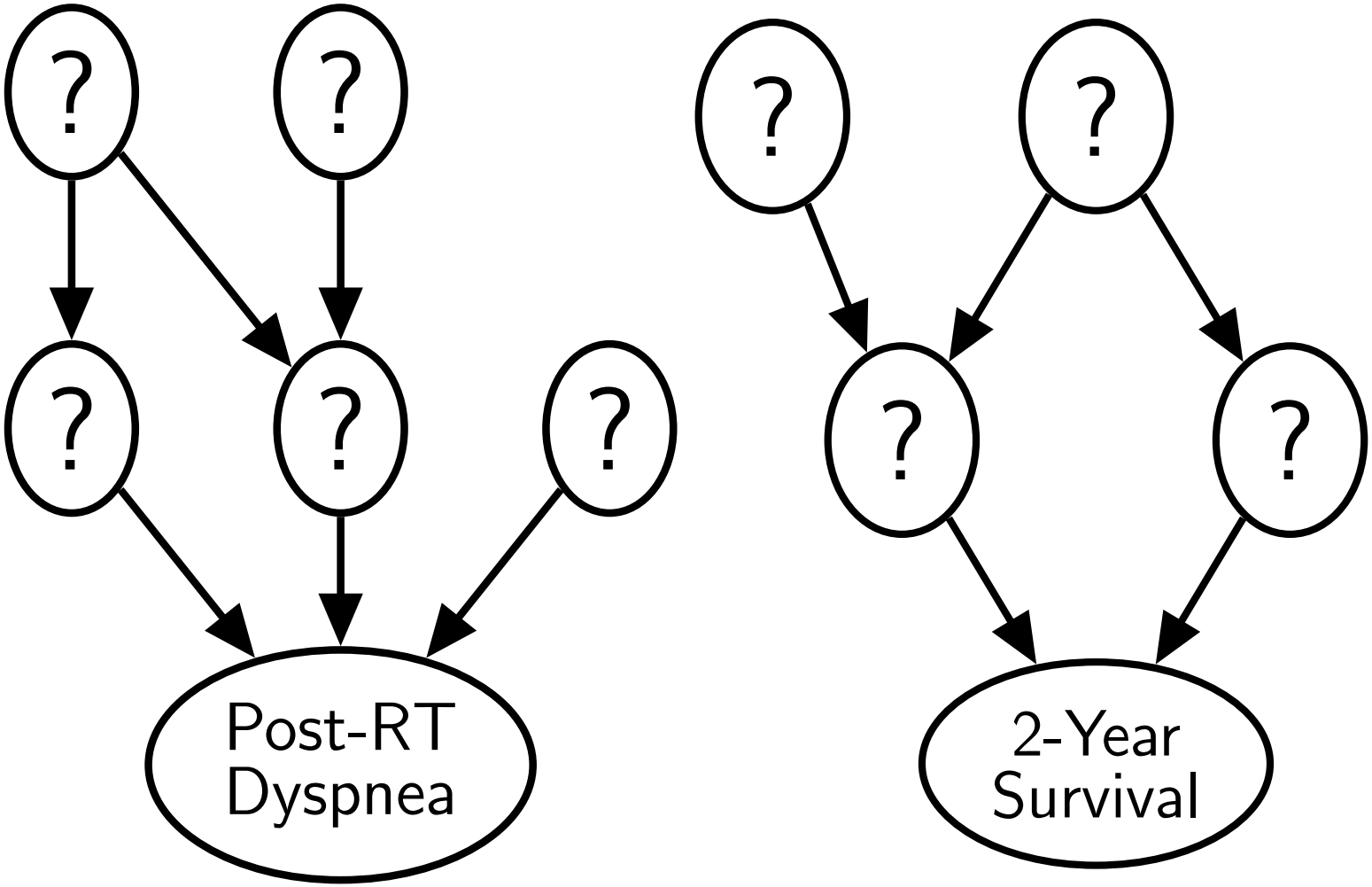
Prob(B. Dysp.|Smoking)

	0	**1**
0	0.6	0.3
1	0.2	0.3
2	0.1	0.2
3	0.1	0.2

Prob(Dysp.|B. Dysp.)

	0	**1**	**2**	**3**	**4**
$< \mathbf{2}$	0.9	0.8	0.3	0.2	0.05
$\geq \mathbf{2}$	0.1	0.2	0.7	0.8	0.95

Target Nodes



Patient Data

2-Year Survival (452 patients)	Dyspnea ≥ 2 (994 patients)
(Chemo-)radiotherapy	(Chemo-)radiotherapy
Curative treatment	Curative treatment
Primary tumor	Primary tumor
NSCLC	NSCLC & SCLC
No surgery	With & without surgery
	Baseline Dyspnea < 2

Missing data is commonplace!

Stratified data split

80% training
20% validation

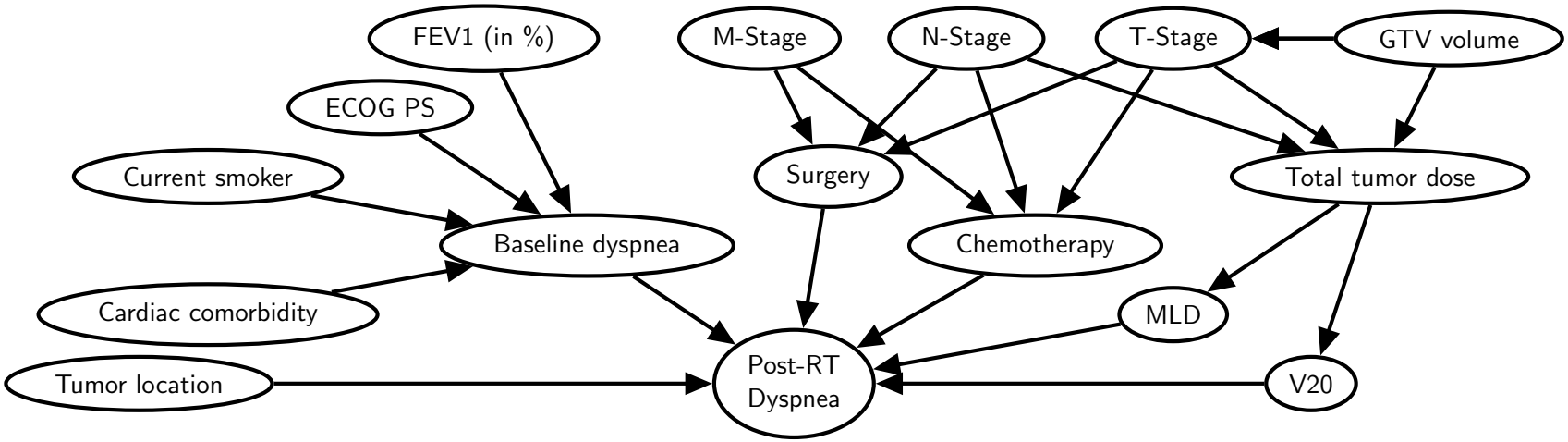
Patient Data

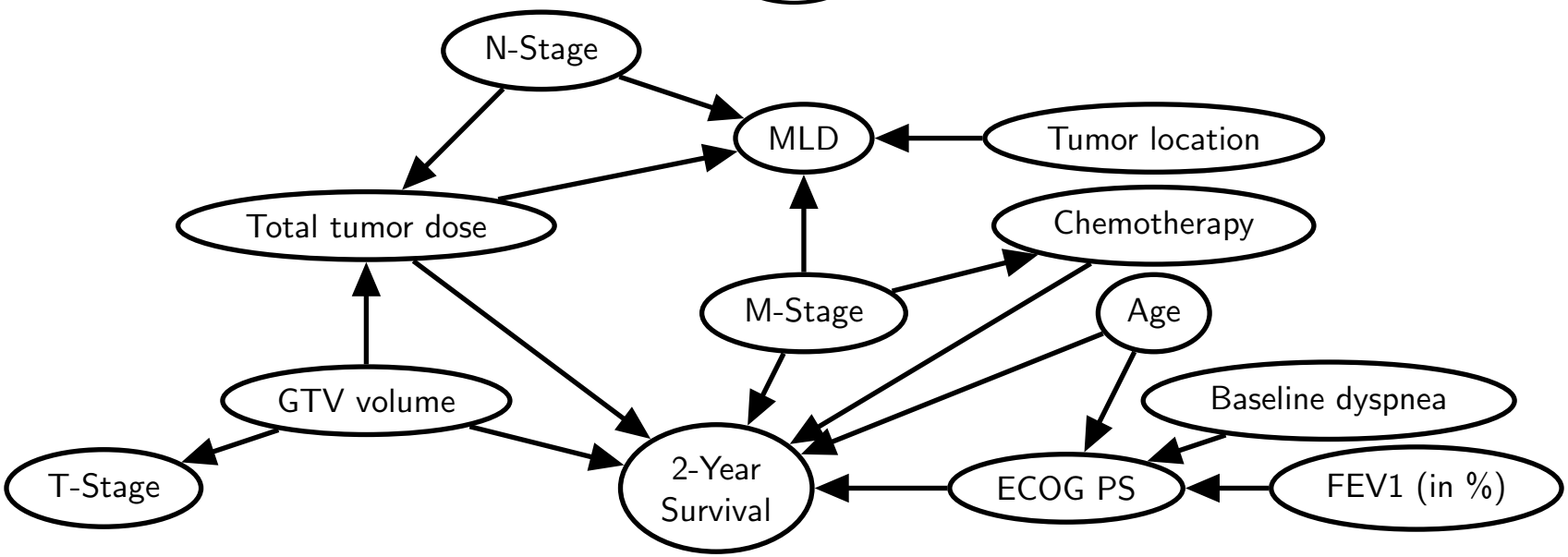
Demographics	Clinical Variables	Disease Variables	Treatment Variables
Age	ECOG PS	Histology	Surgery
Gender	FEV1 (in %)	Tumor Location	Chemotherapy
Smoking Status	Baseline Dyspnea	T-Stage	Modality
		N-Stage	Treatment Time
		M-Stage	Mean Lung Dose
		GTV Volume	$V_{20\%}$
			Tumor Dose

"How do you model dependencies between clinical variables?"

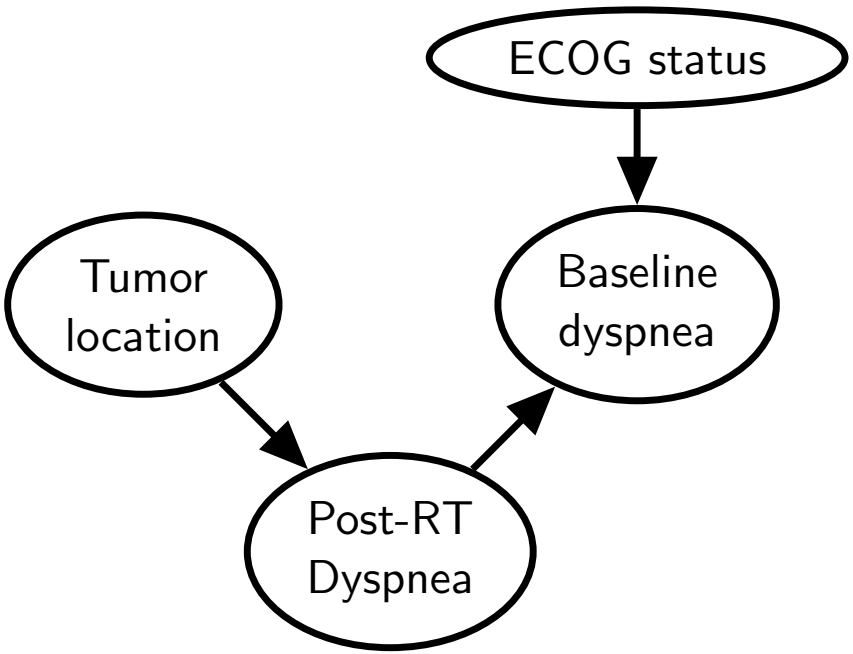
- 10 experts from 5 countries
  - 7 radiation oncologists
  - 3 researchers

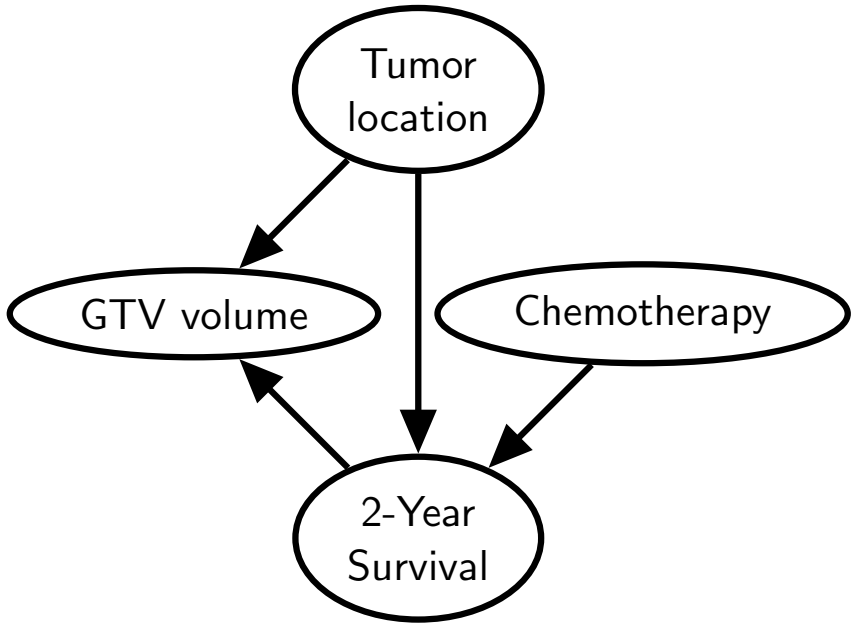
Expert Answers





Algorithm Output





Post-RT Dyspnea ≥ 2

		AUC
Expert	1	0.58
	2	0.61
	3	0.49
	4	0.59
	5	0.65
	6	0.69
	7	0.57
Alg		0.52

[1] Confidence intervals based on 1000 bootstraps

Post-RT Dyspnea ≥ 2

		AUC	AUC 95% CI[1]
Expert	1	0.58	[0.42,0.73]
	2	0.61	[0.43,0.77]
	3	0.49	[0.32,0.65]
	4	0.59	[0.45,0.73]
	5	0.65	[0.5,0.8]
	6	0.69	[0.56,0.83]
	7	0.57	[0.43,0.7]
Alg		0.52	[0.38,0.66]

[1] Confidence intervals based on 1000 bootstraps

Post-RT Dyspnea ≥ 2

		AUC	AUC 95% CI[1]	AUC-AUC$_{Alg}$ 95% CI[1]
Expert	1	0.58	[0.42,0.73]	[-0.07,0.22]
	2	0.61	[0.43,0.77]	[-0.14,0.32]
	3	0.49	[0.32,0.65]	[-0.19,0.15]
	4	0.59	[0.45,0.73]	[-0.06,0.22]
	5	0.65	[0.5,0.8]	[-0.05,0.32]
	6	0.69	[0.56,0.83]	[-0.03,0.39]
	7	0.57	[0.43,0.7]	[-0.08,0.2]
Alg		0.52	[0.38,0.66]	[0,0]

[1] Confidence intervals based on 1000 bootstraps

Post-RT Dyspnea ≥ 2

		AUC	AUC 95% CI[1]	AUC-AUC$_{Alg}$ 95% CI[1]	# Arcs	# Dep.
Expert	1	0.58	[0.42,0.73]	[-0.07,0.22]	30	10
	2	0.61	[0.43,0.77]	[-0.14,0.32]	9	9
	3	0.49	[0.32,0.65]	[-0.19,0.15]	23	7
	4	0.59	[0.45,0.73]	[-0.06,0.22]	22	6
	5	0.65	[0.5,0.8]	[-0.05,0.32]	20	5
	6	0.69	[0.56,0.83]	[-0.03,0.39]	14	5
	7	0.57	[0.43,0.7]	[-0.08,0.2]	7	2
Alg		0.52	[0.38,0.66]	[0,0]	6	2

[1] Confidence intervals based on 1000 bootstraps

2-Year Survival

		AUC
Expert	1	0.59
	2	0.65
	3	0.69
	4	0.56
	5	0.53
	6	0.64
	7	0.68
Alg		0.72

[1] Confidence intervals based on 1000 bootstraps

2-Year Survival

		AUC	AUC 95% CI[1]
Expert	1	0.59	[0.48,0.70]
	2	0.65	[0.54,0.76]
	3	0.69	[0.58,0.80]
	4	0.56	[0.44,0.68]
	5	0.53	[0.40,0.65]
	6	0.64	[0.53,0.75]
	7	0.68	[0.57,0.82]
Alg		0.72	[0.60,0.82]

[1] Confidence intervals based on 1000 bootstraps

2-Year Survival

		AUC	AUC 95% CI[1]	AUC-AUC$_{Alg}$ 95% CI[1]
Expert	1	0.59	[0.48,0.70]	[-0.27,0.01]
	2	0.65	[0.54,0.76]	[-0.21,0.07]
	3	0.69	[0.58,0.80]	[-0.13,0.10]
	4	0.56	[0.44,0.68]	[-0.32,-0.01]
	5	0.53	[0.40,0.65]	[-0.36,-0.01]
	6	0.64	[0.53,0.75]	[-0.21,0.05]
	7	0.68	[0.57,0.82]	[-0.19,0.12]
Alg		0.72	[0.60,0.82]	[0,0]

[1] Confidence intervals based on 1000 bootstraps

2-Year Survival

		AUC	AUC 95% CI[1]	AUC-AUC$_{Alg}$ 95% CI[1]	# Arcs	# Dep.
Expert	1	0.59	[0.48,0.70]	[-0.27,0.01]	19	10
	2	0.65	[0.54,0.76]	[-0.21,0.07]	15	7
	3	0.69	[0.58,0.80]	[-0.13,0.10]	17	6
	4	0.56	[0.44,0.68]	[-0.32,-0.01]	23	8
	5	0.53	[0.40,0.65]	[-0.36,-0.01]	13	7
	6	0.64	[0.53,0.75]	[-0.21,0.05]	16	6
	7	0.68	[0.57,0.82]	[-0.19,0.12]	20	11
Alg		0.72	[0.60,0.82]	[0,0]	4	3

[1] Confidence intervals based on 1000 bootstraps

Summary

- Experts & algorithm fail to predict dyspnea

  $\Rightarrow$ **New biomarkers needed**

Summary

- Experts & algorithm fail to predict dyspnea

  $\Rightarrow$ **New biomarkers needed**

- Algorithm predicts survival (slightly) better

Summary

- Experts & algorithm fail to predict dyspnea
  $\Rightarrow$ **New biomarkers needed**
- Algorithm predicts survival (slightly) better
- Simpler models yield the same or better results

Summary

- Experts & algorithm fail to predict dyspnea

  $\Rightarrow$ **New biomarkers needed**

- Algorithm predicts survival (slightly) better

- Simpler models yield the same or better results

- Experts' complex systems not reflected in data
  - Overestimated effects?
  - Too complex for available data

  $\Rightarrow$ **Share data**

Thank you for your attention!

Timo M. Deist
timo.deist@maastro.nl

Acknowledgments

MAASTRO Clinic

Arthur Jochems
Cary Oberije
Bart Reymen
Rinus Wanders
Wouter van Elmpt
Andre Dekker
Philippe Lambin

Universiteit Ghent

Katrien Vandecasteele
Yolande Lievens

Karolinska Institutet

Karin Lindberg

KU Leuven

Dirk De Ruysscher

The Christie (NHS)

Corinne Faivre-Finn

University of New South Wales

Shalini Vinod