The role of analytics in improving the efficiency of cancer treatment facilities

LOUIS-MARTIN ROUSSEAU

CANADA RESEARCH CHAIR IN
HEATLHCARE ANALYTICS AND LOGISTICS
HANALOG

POLYTECHNIQUE
MONTRÉAL

Healthcare systems

One person is diagnosed with cancer every 3 minutes in Canada, 20 seconds in USA.

One person dies from cancer every 7 minutes in Canada, 1 minute in USA.

First cause of mortality in Canada (30%):

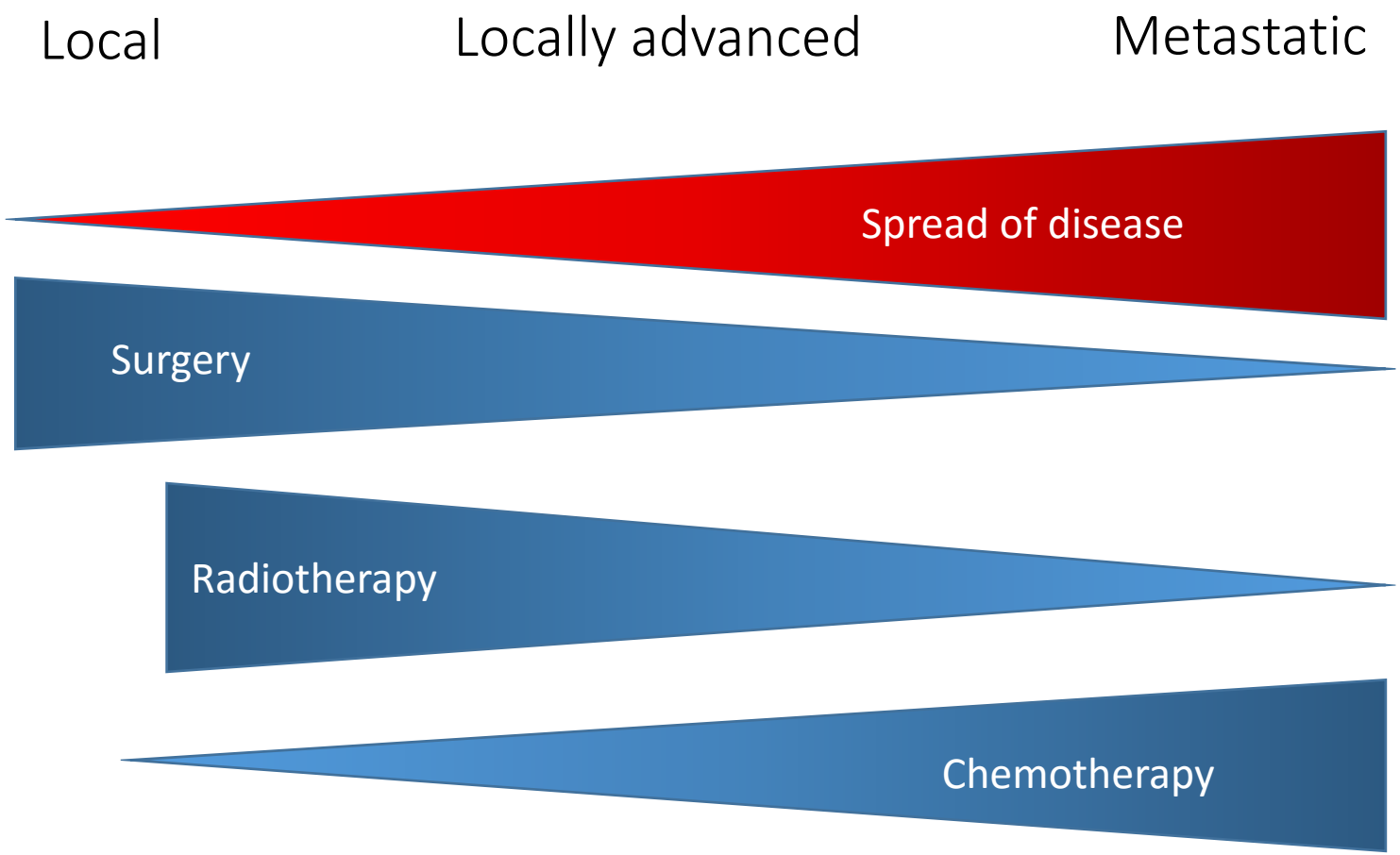
- 45% of Canadian will develop cancer
- 5 year survivability 66%

Ever increasing of new cancer cases:

- 12% within 4 years
- Aging of population;
- Demographic growth.

How to treat all these patients while keeping excellent care ?

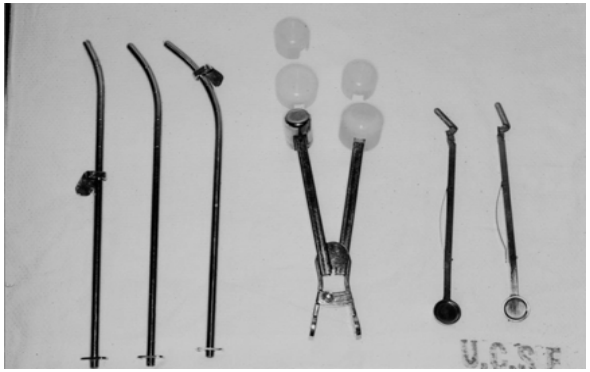
What are your treatment options ?

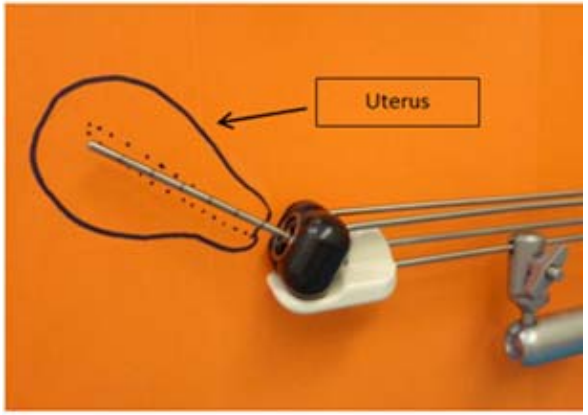


About 50% of cancer patients will receive radiotherapy

Tools

Internal





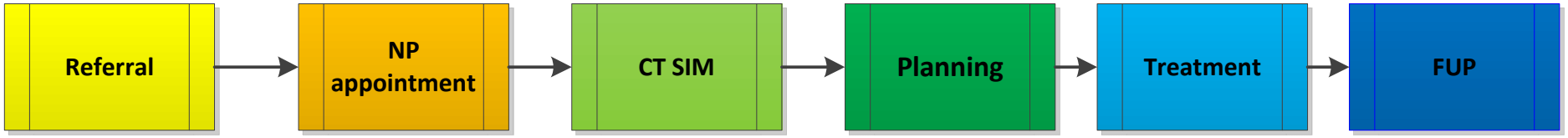
External

Vickers 6 Prototype Newcastle-on-Tyne 1960

Teams

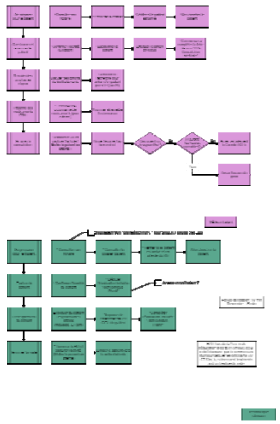
	Chemotherapy	Radiotherapy
Prescribes	Oncologist	Radiation Oncologist
Prepares	Pharmacist	Physicist
Delivers	Nurse	Therapist

Care Trajectory

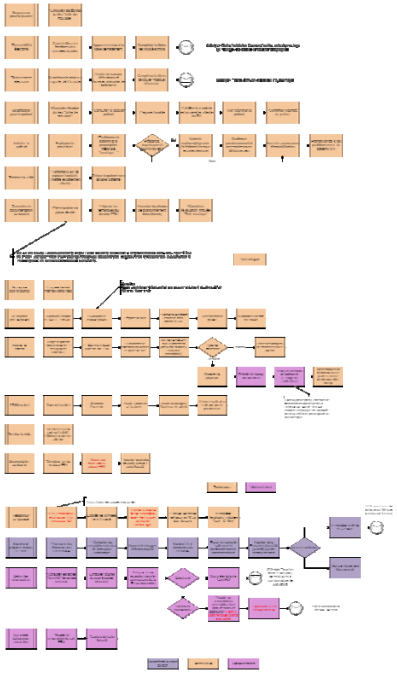


Care Trajectory in details

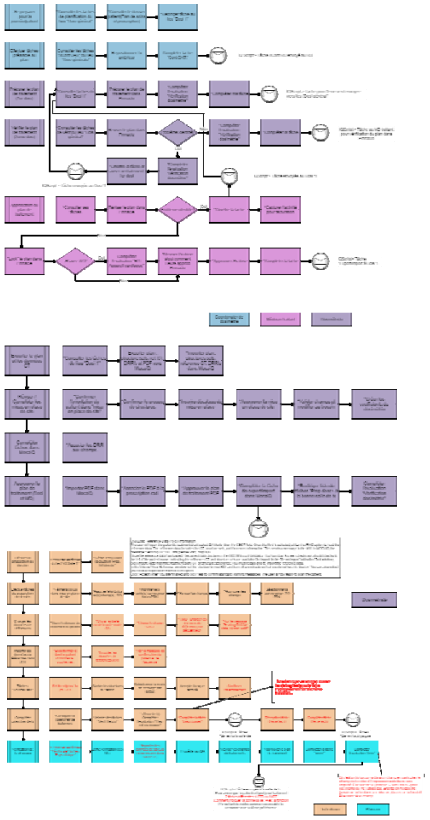
NP appointment



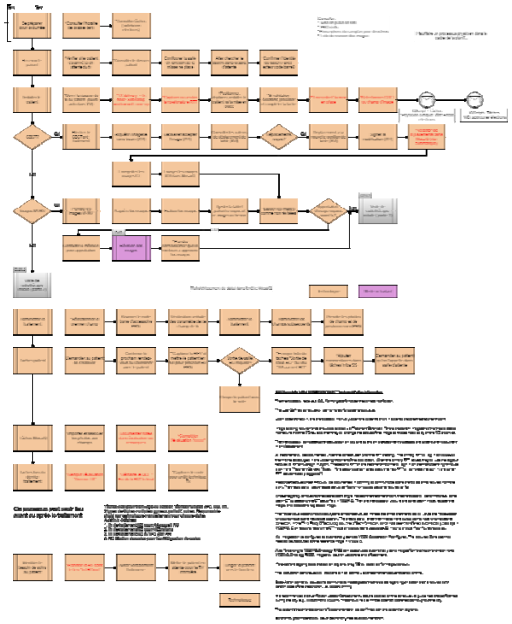
CT SIM



Planning



Treatment



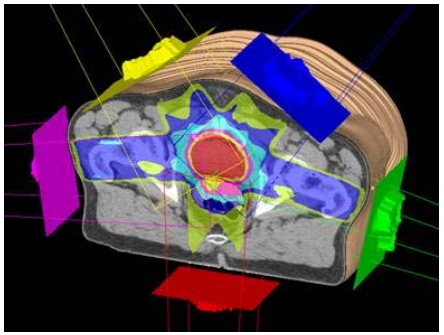
Important steps

Simulation:

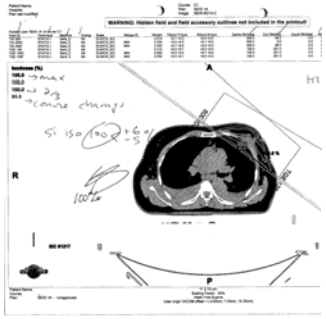
- Uses: CT, MRI, PET-CT
- Used for treatment planning purposes
- 3D model of the human body

Treatment Planning

- Calculates radiation deposition in the human body
- Multi-criteria optimization solver
- Server farm, GPU calculations, etc.



Plan approval



Linear accelerator

- mm accuracy
- 100x more powerful than a radiology X-ray

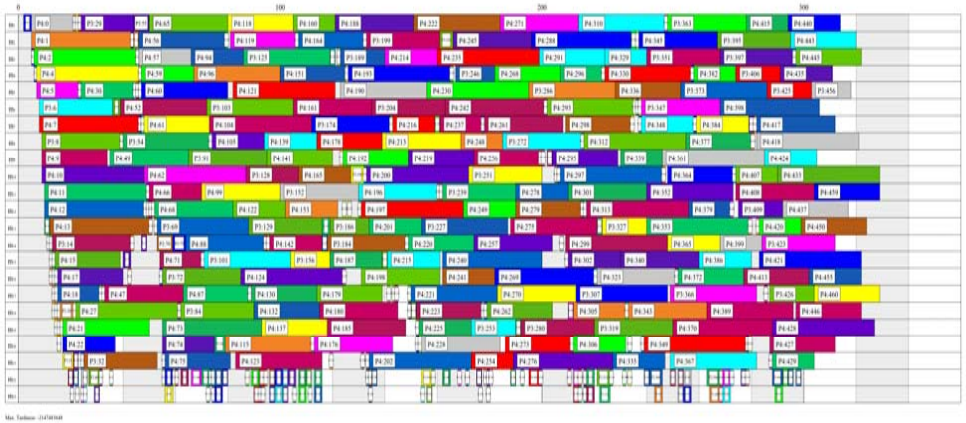
(Q1) when to book a patient ?

Patient arrives

how much time ?

Patient is treated

Considering existing calendar…



… and patient priorities

Palliative	Curative 1	Curative 2
< 3 days	< 14 days	< 28 days

Different possible approaches

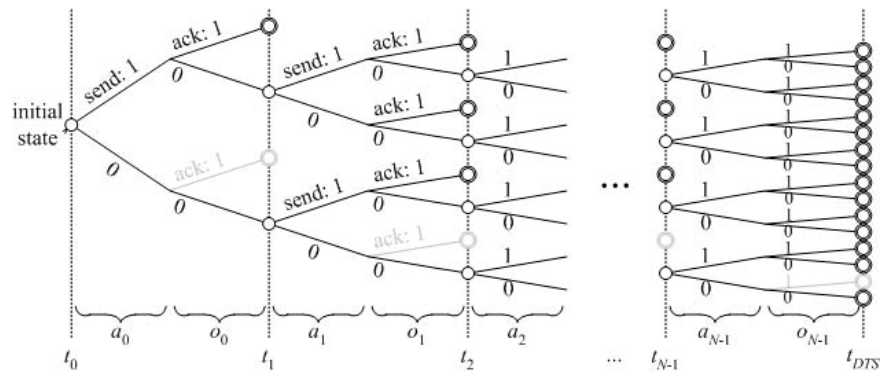
Stochastic Optimization

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & g(x) = c^T x + E[Q(x,\xi)] \\ \text{subject to} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

$$\Longleftrightarrow$$

$$\begin{aligned} \min_{y \in \mathbb{R}^m} \quad & q(\xi)^T y \\ \text{subject to} \quad & T(\xi)x + W(\xi)y = h(\xi) \\ & y \geq 0 \end{aligned}$$

Markov Decision Process



Online Optimization

Online Stochastic Combinatorial Opt.

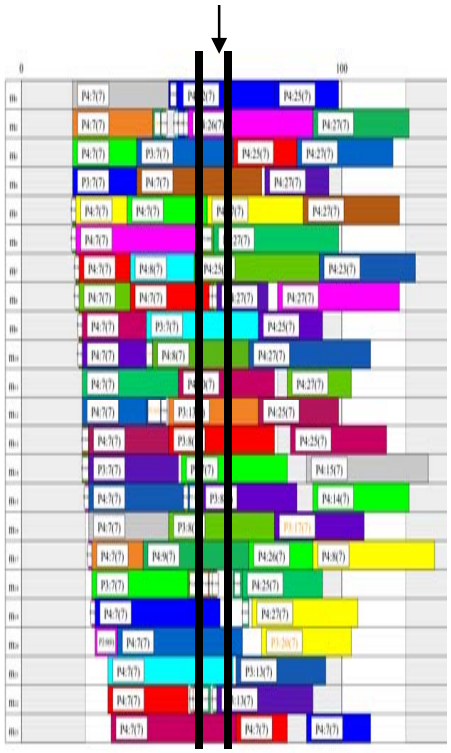
For each new request:

- Build a set of scenarios;
- Compute a solution for each scenario
- Heuristically, choose a response for the request based on all solutions.

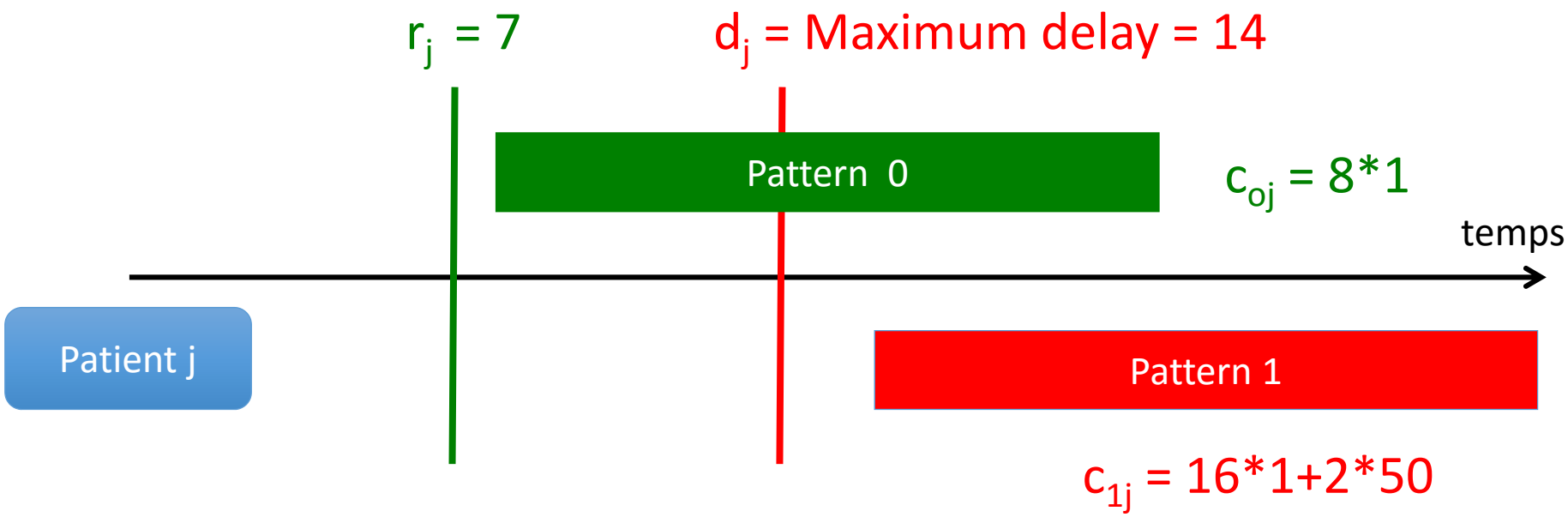
OSCO in context RT cancer patient booking

- Online stochastic combinatorial optimization:
  1. For each solution, we compute :
     1. A utilization cost (by day and by linac) for a time slot;
  2. We choose the appointment of minimum cost:
     1. Waiting time cost (depending of the priority) ;
     2. Expected utilization cost.
- Booking model -> Dantzig-Wolfe decomposition;
- Uncertainties -> Benders decomposition.

Utilization cost



Structure of the model



Treatment planning fix to 7 days … for now

Stochastic Programming Model

$$\min \sum_{i \in S_j} c_{ij} x_{ij} \mid \mathbb{E}_{\omega \in \Omega_j}\Big[\sum_{l \in \mathcal{P}^\omega} \sum_{i \in S_l} c_{il} y_{il}^\omega \mid \sum_{k \in H} \sum_{m \in M} c^o z_{mk}^\omega\Big]$$

subject to:

$$\sum_{i \in S_j} x_{ij} = 1$$

$$\sum_{i \in S_l} y_{il}^\omega = 1, \qquad \forall \omega \subset \Omega_j, \forall l \subset \mathcal{P}^\omega$$

$$\boxed{\sum_{i \in S_j} a_{ijk}^m x_{ij} + \sum_{l \in \mathcal{P}^\omega} \sum_{i \in S_l} a_{ilk}^m y_{il}^\omega \leq F_k^m + z_{mk}^\omega, \qquad \forall m \in M, \forall k \in H, \forall \omega \in \Omega_j}$$

$$\mathbb{1}_{\mathcal{P}_p}(j) \sum_{i \in S_j} a_{ijk}^m x_{ij} + \sum_{l \in \mathcal{P}_p^\omega} \sum_{i \in S_l} a_{ilk}^m y_{il}^\omega \geq z_{mk}^\omega, \qquad \forall m \in M, \forall k \in H, \forall \omega \in \Omega_j$$

$$\sum_{k=b}^{b+4} z_{mk}^\omega \leq O_{week}, \qquad \forall m \in M, \forall b \in \mathcal{B}, \forall \omega \in \Omega_j$$

$$z_{mk}^\omega \in [0, O_{day}], \qquad \forall m \in M, \forall k \in H, \forall \omega \in \Omega_j$$

$$x_{ij} \in \{0,1\}, \qquad \forall i \in S_j$$

$$y_{il}^\omega \in \{0,1\}, \qquad \forall l \in \mathcal{P}^\omega, \forall i \in S_l, \forall \omega \in \Omega_j$$

Approximated utilization cost of a given initial treatment time slot

= dual variable of this constraint

Choose greedily the pattern With best reduced cost

Pattern 1

Pattern 0

Initial Results

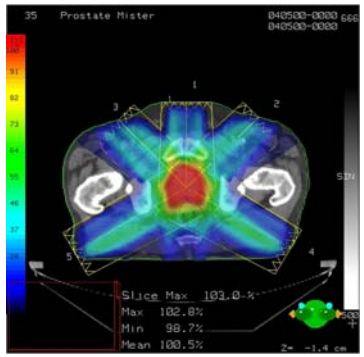
	Due date violations			Average waiting time			Utilization	Overtime
	>3	>14	>28	Palliative	Curative 1	Curative 2		
CICL	14	16	0	2,07	14,38	12,98	88,3%	44
OSCO - 1	9	6	0	1,05	10,57	15,98	88,0%	6

CICL real data:

- 170 patients ;
- 120 days;
- 2 linacs with 23 slots.

Legrain A, MA Fortin, Lahrichi N, Rousseau L-M (2015) “Online Stochastic Optimization of Radiotherapy Patient Scheduling”, *Healthcare Management Science*, 18, 110-123.

(Q2) How much for treatment preparation ?

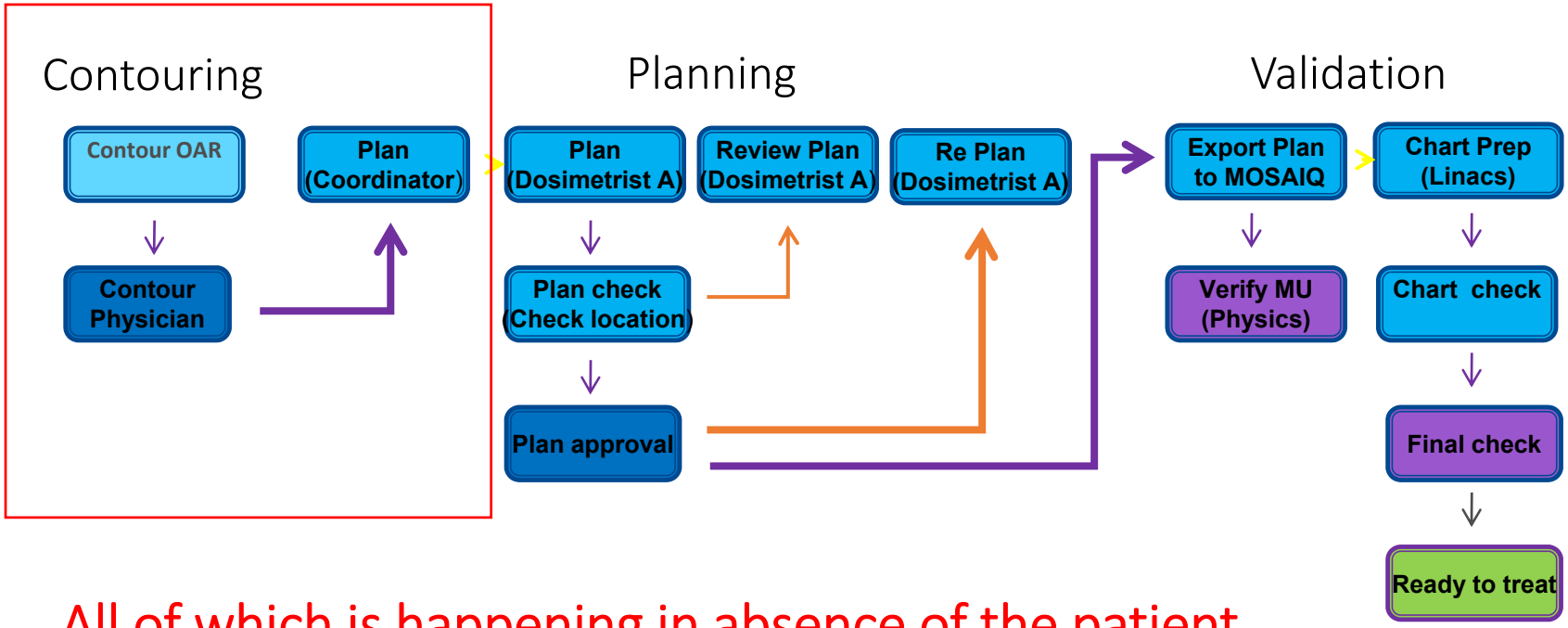


Appointment booking

Unknown dosimetry duration

Preparation completed

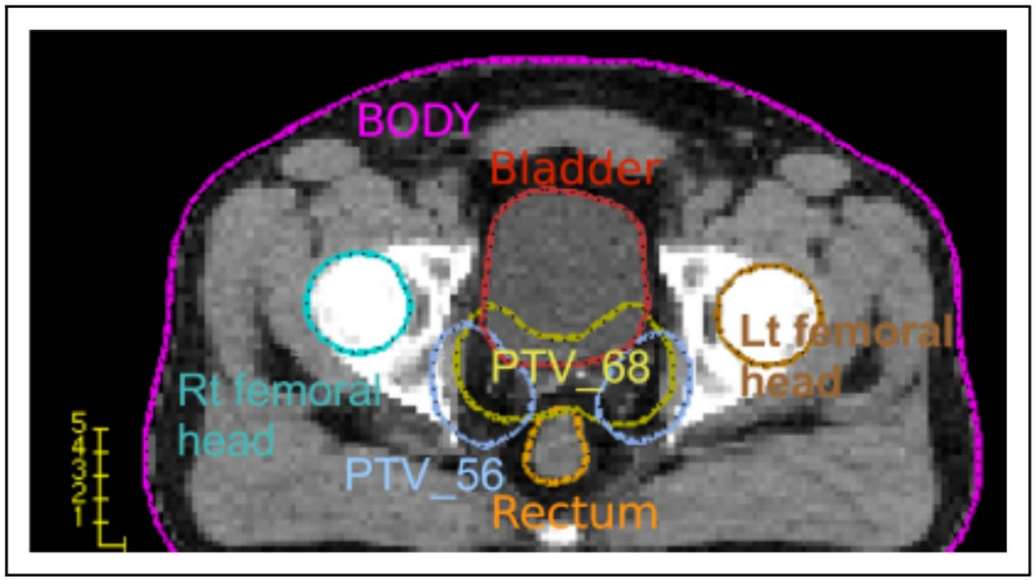
PreparationTasks



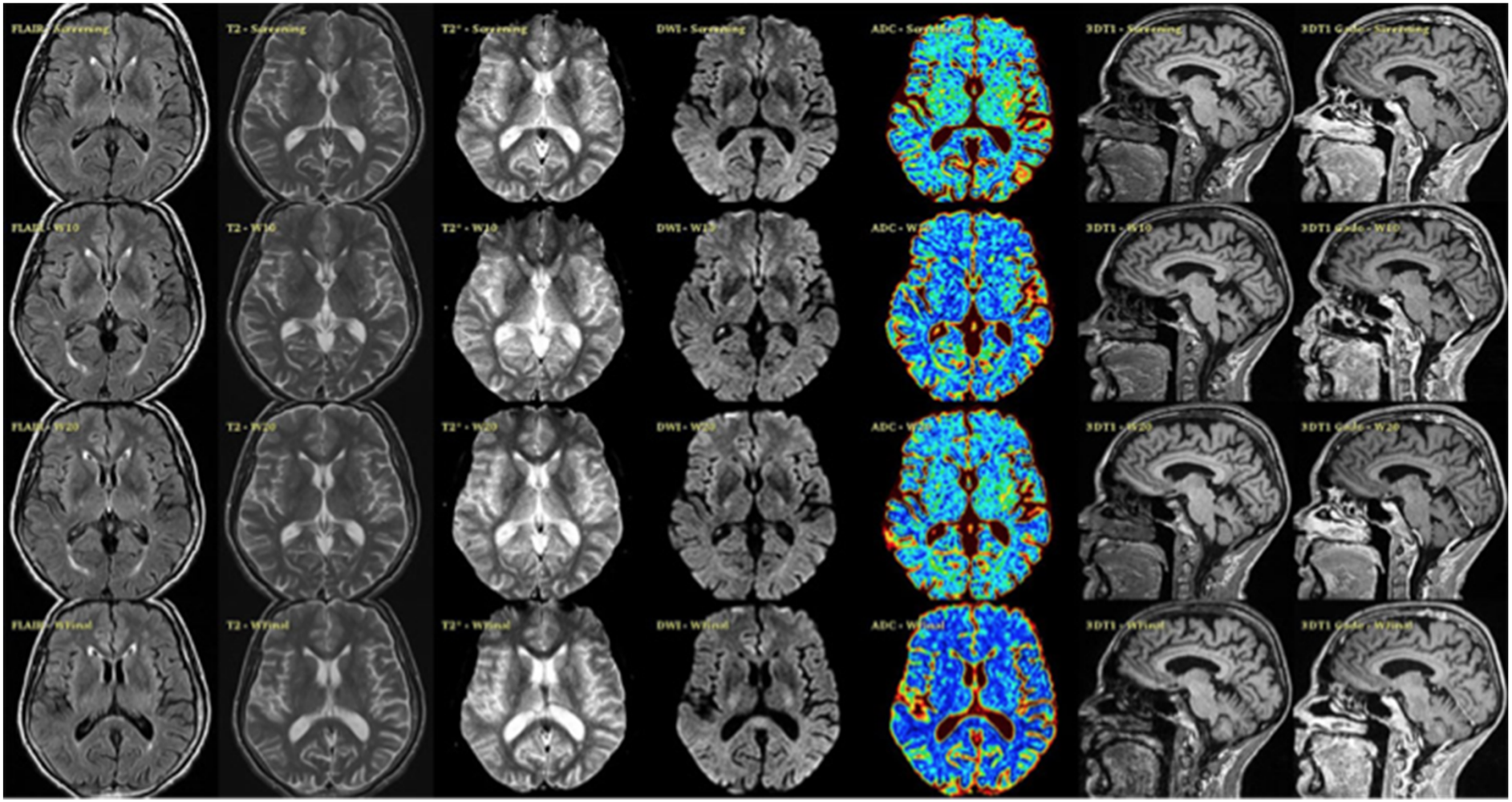
All of which is happening in absence of the patient

The contouring problem

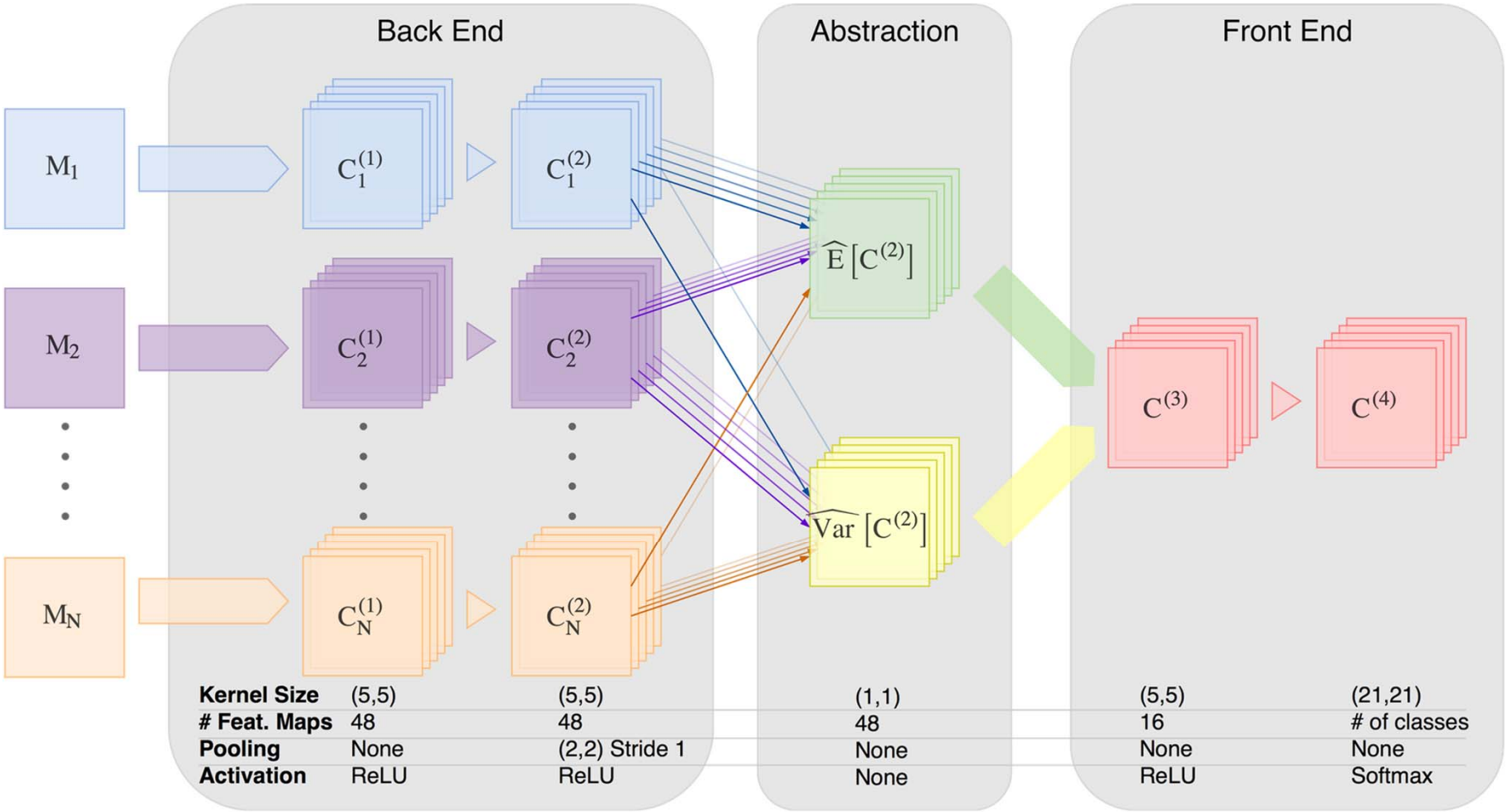
- Radio Oncologists look at many images, patient history, medical file, etc.
- They build a “mental model” where is tumor and how it should be attacked.
- Then they project this “model” on a screen



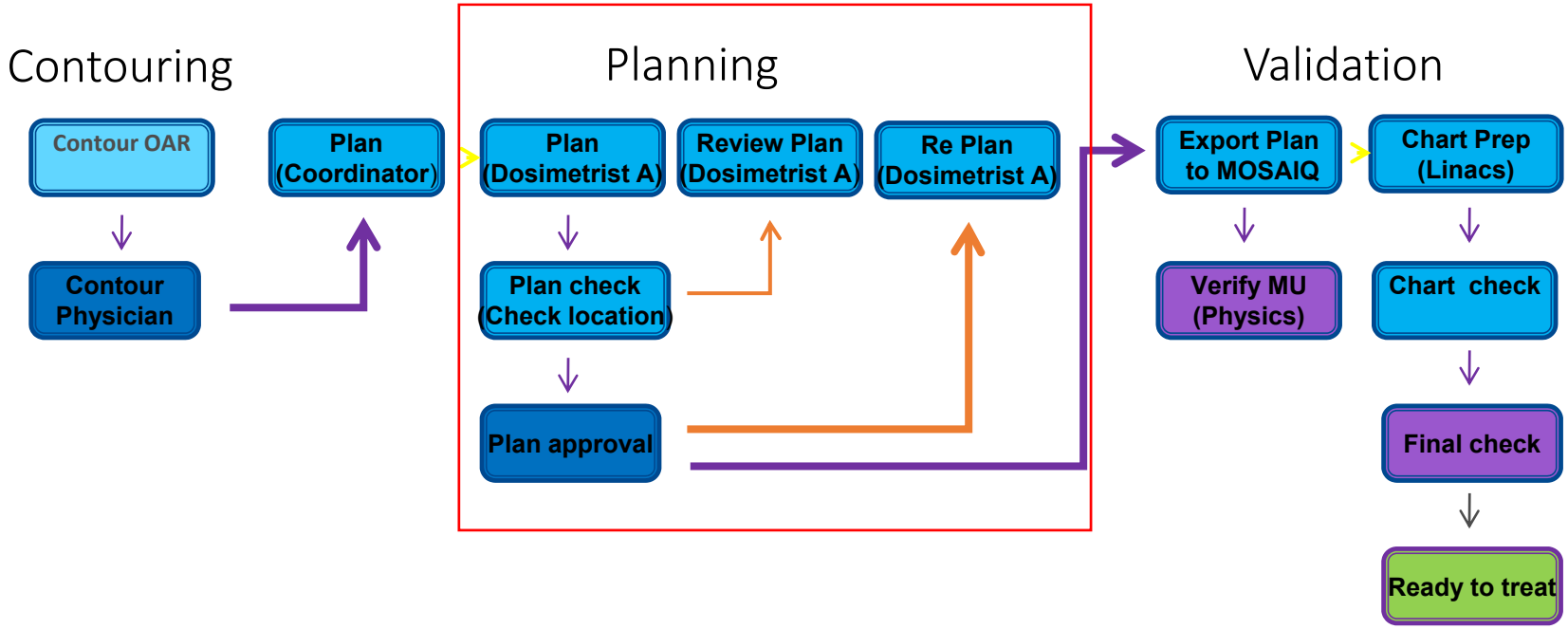
Medical Imaging is a rich source of Data



Deep Neural Network



Preparation Tasks



Best feasible appointment

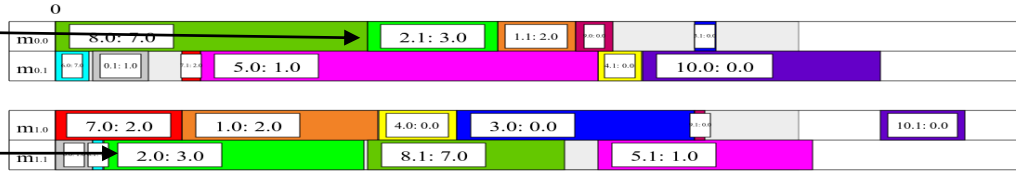
Online stochastic optimization:

1. For each solution, we compute :
   1. A utilization cost (by day and by linac) for a slot;

   2. A minimum day to start the treatment.
      1. with fast Genetic Algorithm
      2. with exact Constraint Programming Model

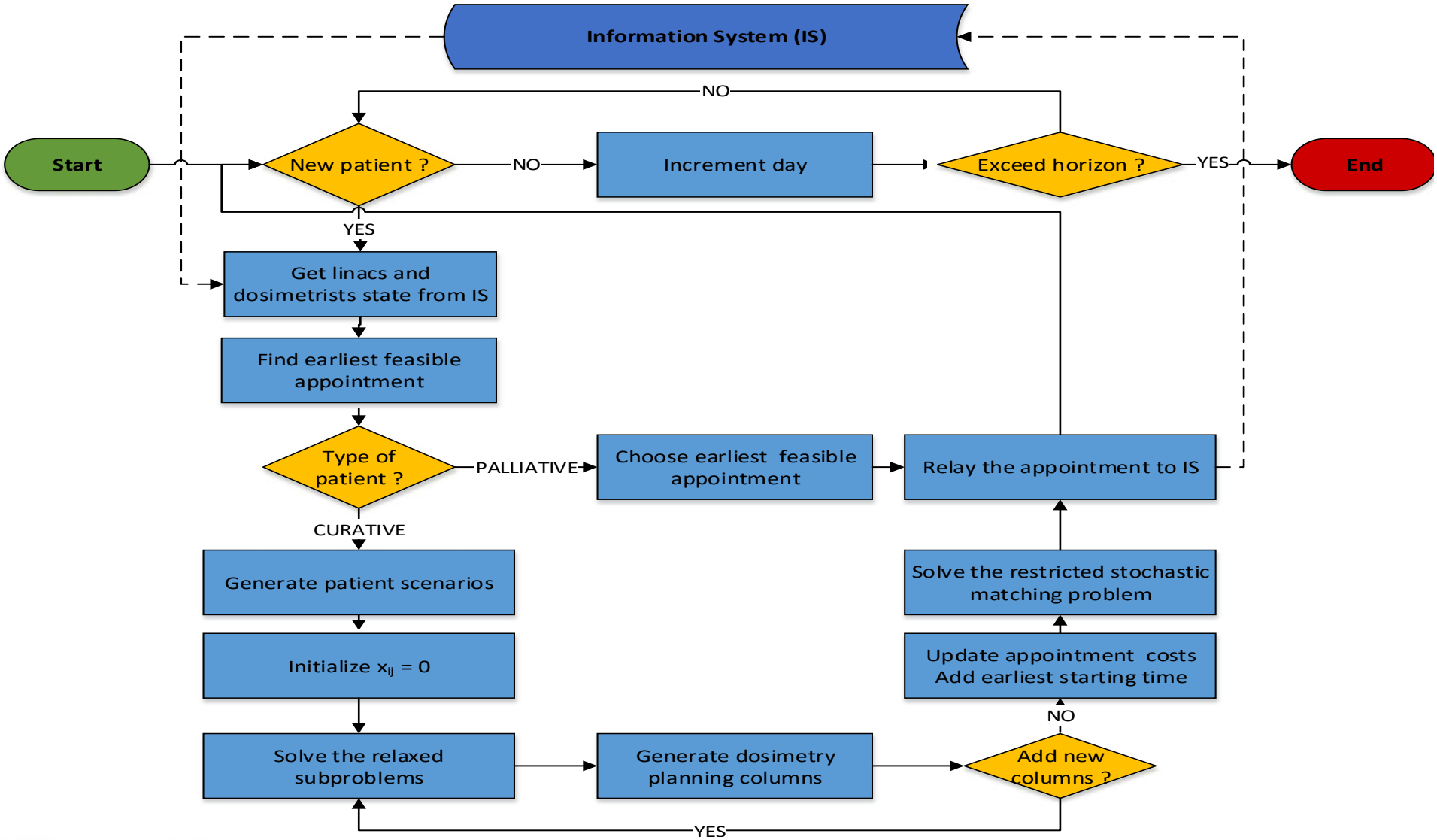
Verification

Preparation



2. We choose the appointment of minimum cost:
   1. Waiting time cost (depending of the priority) ;
   2. Expected utilization cost.

Global Approach



New Results

	Cancellations	Due date (in days)			Average waiting time			Overtime
		>3	>14	>28	>3	>14	>28	
CICL	230	373	104	0	3,45	12,58	12,63	111
OSCO 1	107	335	67	0	1,05	10,57	15,98	19
OSCO 2	1	326	119	0	3,23	14,04	18,43	8

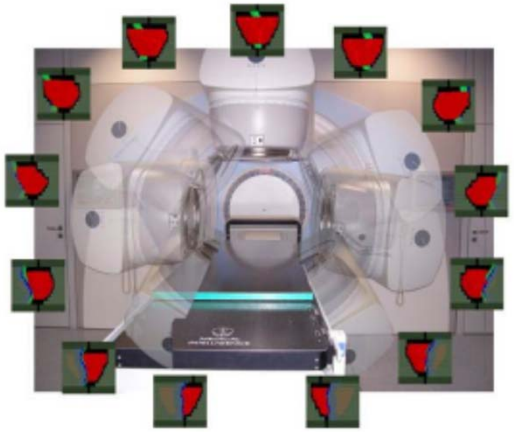
CICL real data:

- 1529 patients ;
- 248 days;
- 4 linacs with 29 slots.

A. Legrain, N. Lahrichi, LM. Rousseau and M. Widmer, Combining Benders and Dantzig-Wolfe decompositions for online stochastic combinatorial optimization, *submitted to IJoC*.

Can treatment be delivered faster ?

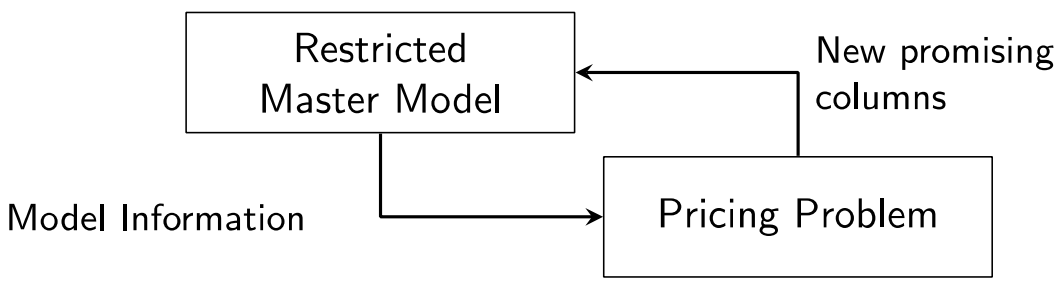
- More details on VMAT



- The VMAT treatment planning problem consists in:
  - Selecting a delivery sequence of collimator shapes
  - Determining the optimal dose rate and rotation speed.
- Our Objective
  - Maximize plan quality
  - Minimize treatment time
  - Fast computation time on “cheap” hardware

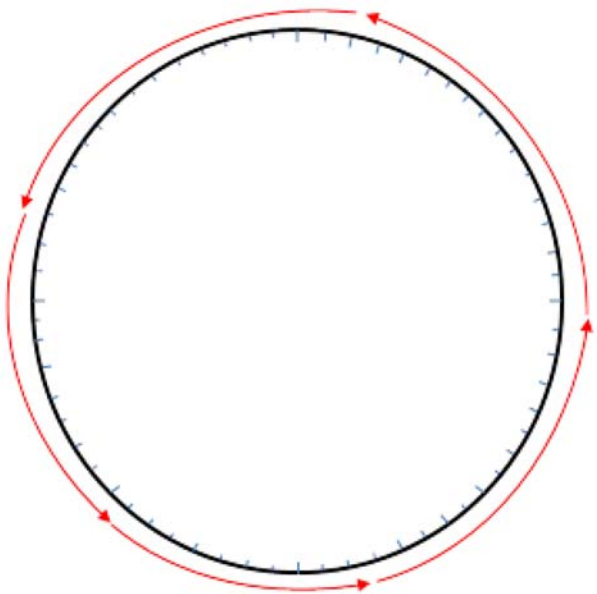
Column Generation Approach

- Highly combinatorial problem:
  - In a small case with a (5 × 10) beam and 100 sectors,
  - there are 7.1×10^{251} apertures shapes.
- Column generation (CG) is a leading optimization technique successful solving large-scale problems
  - Exploits decomposable structures
  - Handles large number of variables



Column Generation Approach

- 360° around the patient is covered by arcs,
- Each arc consists of fixed sectors determining:
  - the aperture shape for including sectors
  - gantry speed
  - dose rate



Master Model: arc and intensity selection

Objective function

- quadratic voxel-based penalty function + delivery time

Constraints

1. calculating the dose deviation from descripted thresholds
2. Each sector should be covered at most by one arc
3. Restricting the change of dose rate between adjacent sectors
4. Restricting the dose rate to the max R
5. The gantry speed at each sector should be enough for leaf motions of the assigned arc
6. Restricting the change of sector time between adjacent sectors
7. Restricting the sector time to lower and upper bounds
8. Restricting the maximum total treatment time.

Master Model: arc and intensity selection

$$\mathbf{GP}: \min \quad \boxed{\mathbf{F}(z) + w\,T_{max} \qquad \text{Weighted quality and time objective}}$$

$$z_j = \sum_{k \in K} \sum_{h \in H_k} D_{jh}(A_h^k)\, y^k\, \rho_h\, t_h \qquad \forall j \in \mathcal{V}$$

$$\sum_{k \in K} a_h^k\, y^k \leq 1 \qquad \forall h \in H$$

$$|\, \rho_{h+1} - \rho_h \,| \leq \Delta_\rho \qquad \forall h = 1, 2, \cdots, |H| - 1$$

$$0 \leq \rho_h \leq R \qquad \forall h \in H$$

$$\sum_{k \in K} \tau_{h,h+1}^k\, y^k \leq t_h \qquad \forall h \in H$$

$$|\, t_{h+1} - t_h \,| \leq \Delta_t \qquad \forall h = 1, 2, \cdots, |H| - 1$$

$$\underline{T} \leq t_h \leq \overline{T} \qquad \forall h \in H$$

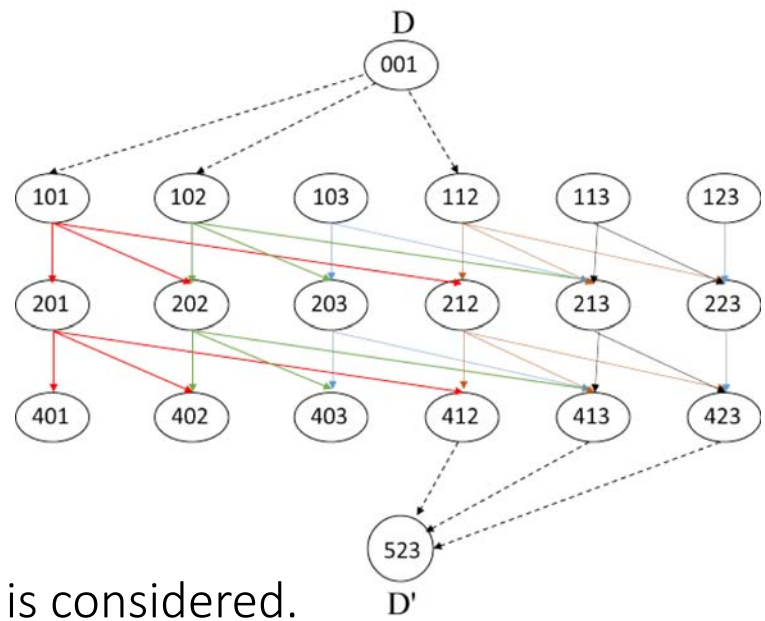
$$\sum_{h \in H} t_h \leq T_{max}$$

$$y^k \in \{0, 1\} \qquad \forall k \in K$$

Subproblem: building new arcs

The situation of each row in each sector is indicated as a node (h, l, r);

- e.g. node $(90, 0, 4)_5$ is the position of leaves of row 5 in sector 90:



Constraints include:

1. Maximum leaf motion constraint is considered.
2. Conflicting trailing and leading leaves are avoided, i.e. $t + 1 \leq r$
3. Cost of nodes and arcs based on the Master Model (dual values)

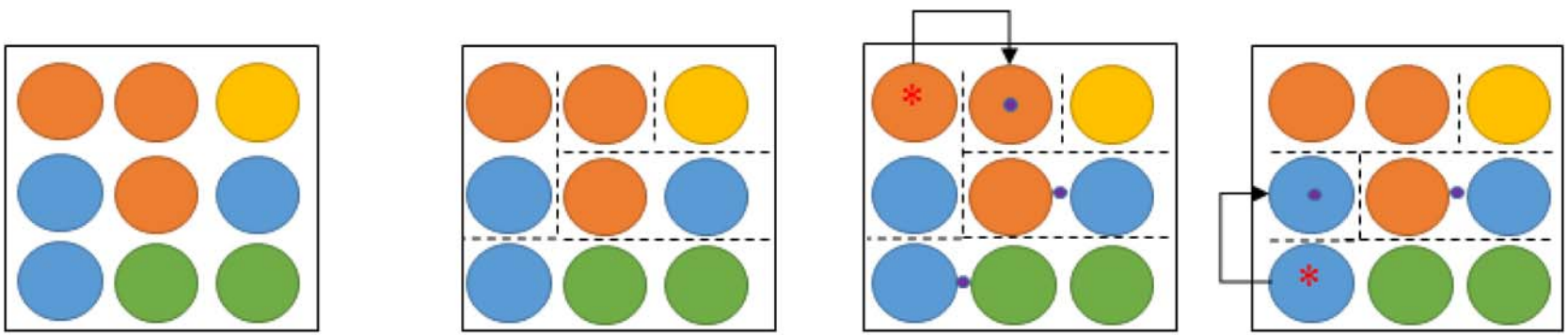
- Polynomial shortest path algorithm easily obtain the best solution.

Reducing problem size

Random down-sampling is a usual approach (Kufer et al. 2003)

We rather propose a method:

- Inspired by K-Means algorithm, the well-known data mining technique
- The goal: Similar neighbor voxels would be considered in a cluster.
- Value of voxel given by full open radiation of all beamlets.
- Each voxel reprensted by a vector in space R^n (n = number of beamlets).



Step 1 Step 2 Step 3

Reducing problem size

- Normal voxels are reduced to 5% and tumor voxels to 10%.
- The progress of solution quality from random sampling to the first iteration was about 78%.
- The progress from the first iteration to the fifth iteration has been only 9%

Voxel aggregation computational results.

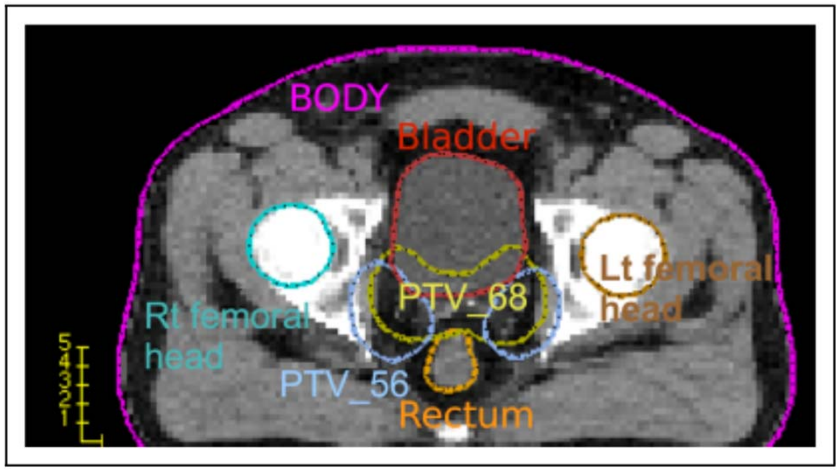
Iter	# Transfer Iter	Time (Sec.)	Avg.Dist
0			54.27054
1	19265	0.952947	11.57258
2	2160	0.74742	10.51782
3	301	0.733188	10.44499
4	40	0.729702	10.44053
5	13	0.738105	10.43969

Experimental evaluation

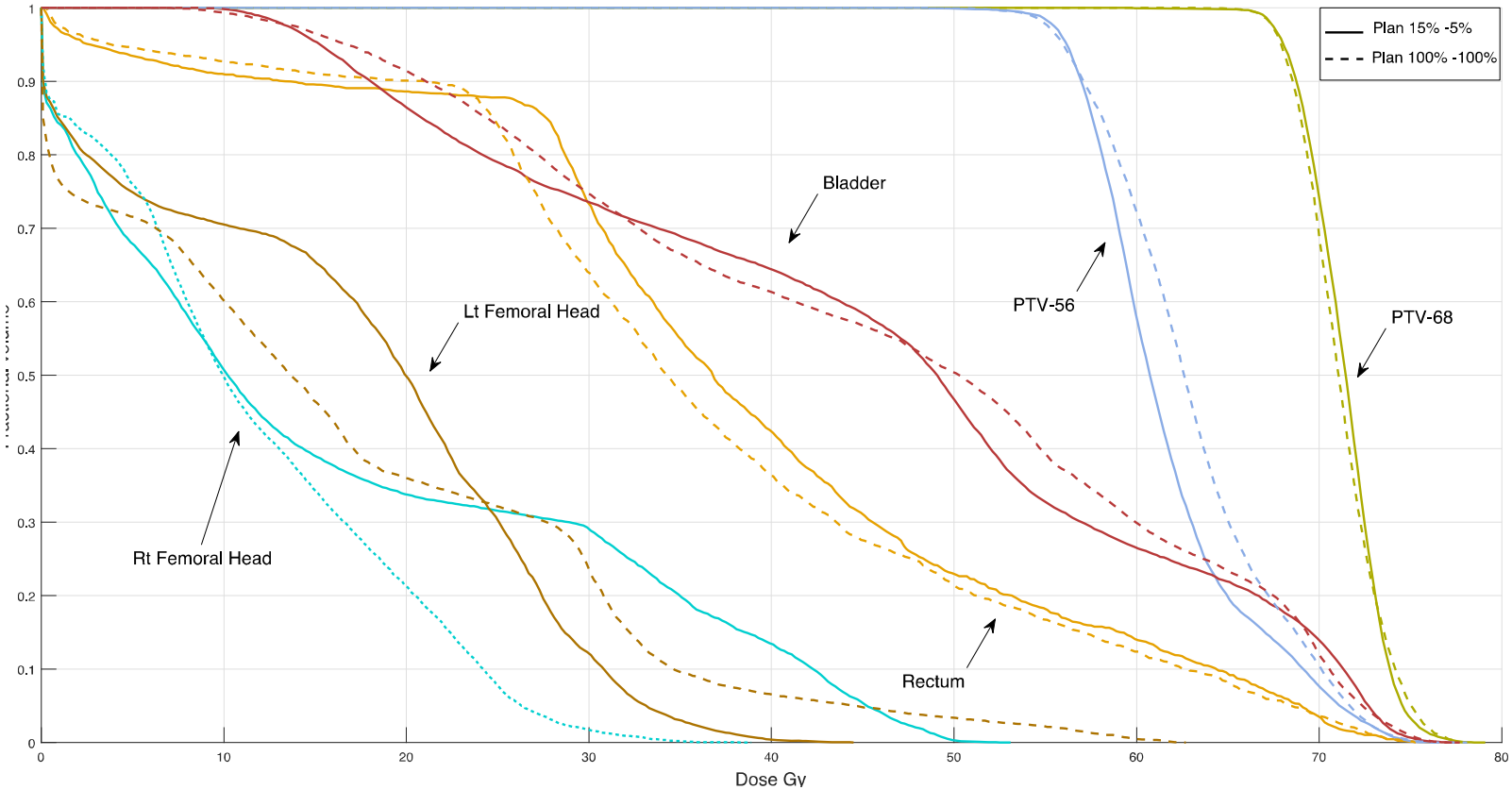
Case Characteristics	
Total # beamlets	25,404
Beamlet size (mm)	1×1
Voxel resolution (mm)	3, 3, 3
# Target voxels	9491
# Body voxels	690,373

Algorithm Parameters	
Max dose rate	600 MU/min
Max leaf speed	3 cm/sec
Max fluence change	2 MU/s
Max time change	2 s
Gantry speed	$[1 \quad 6]^\circ$/sec

- CORT dataset (Craft et al, 2014)
- 180 equispaced sectors
- Algorithm is implemented in C++/CPLEX



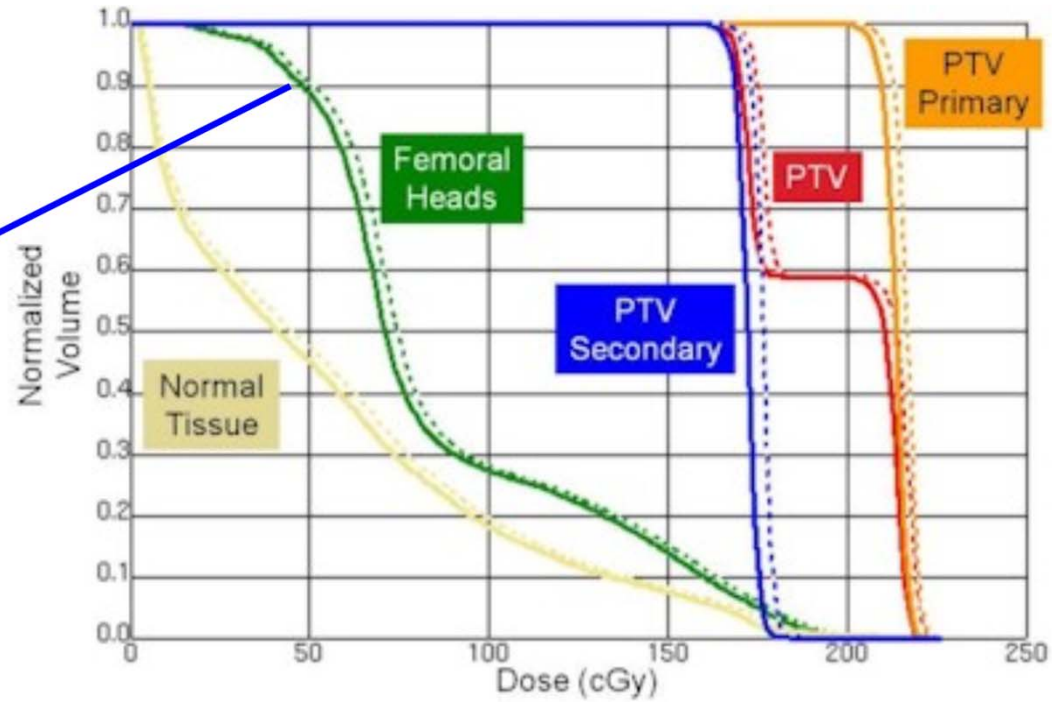
Effect of ML-based aggregation



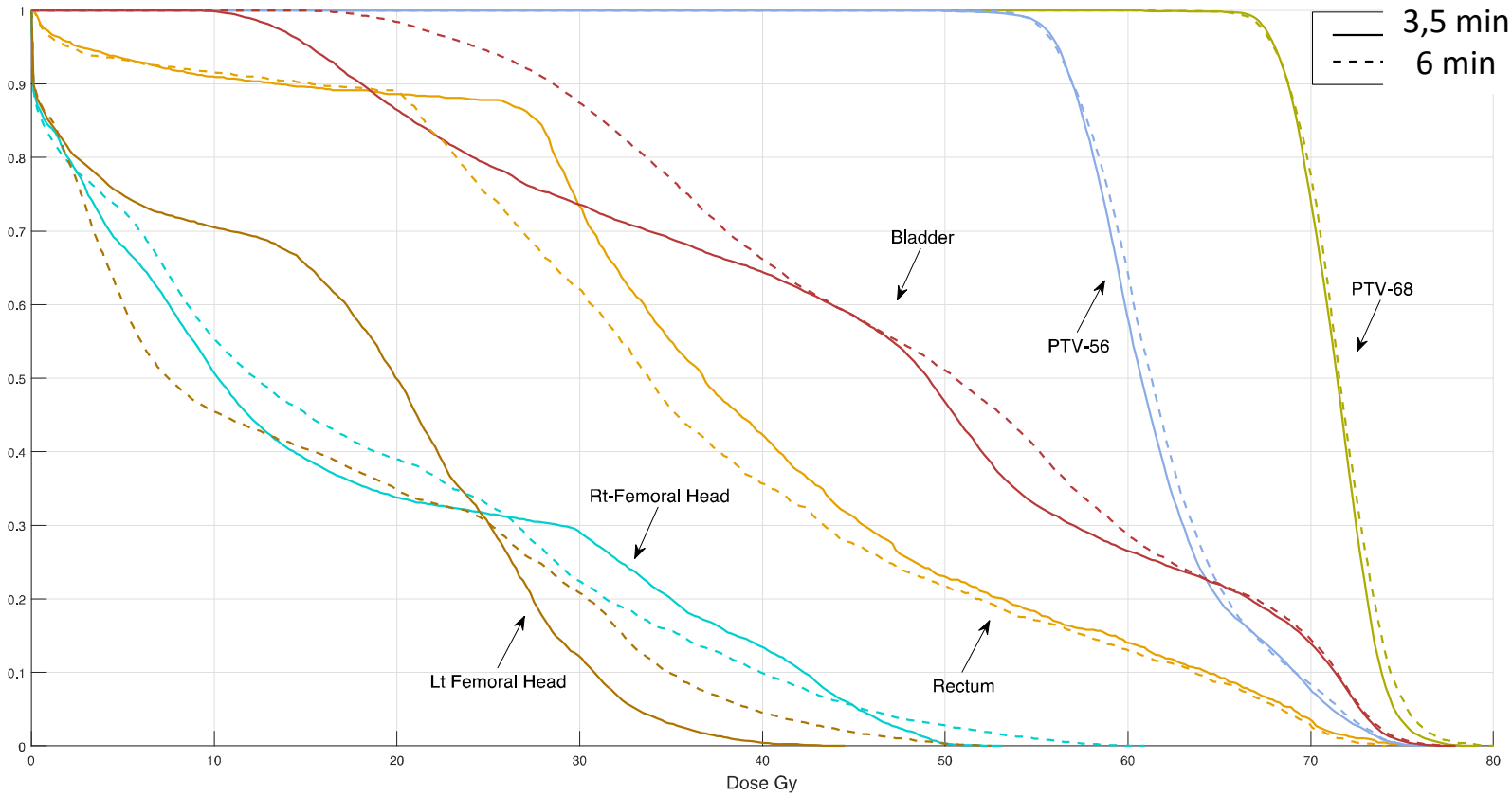
CPU Times
Agg: 4,5 min
Full: 29 min

Dose Volume Histogram

No more than 90 of the healthy organ is allowed to receive 50 Gy or more



Effect of delivery time

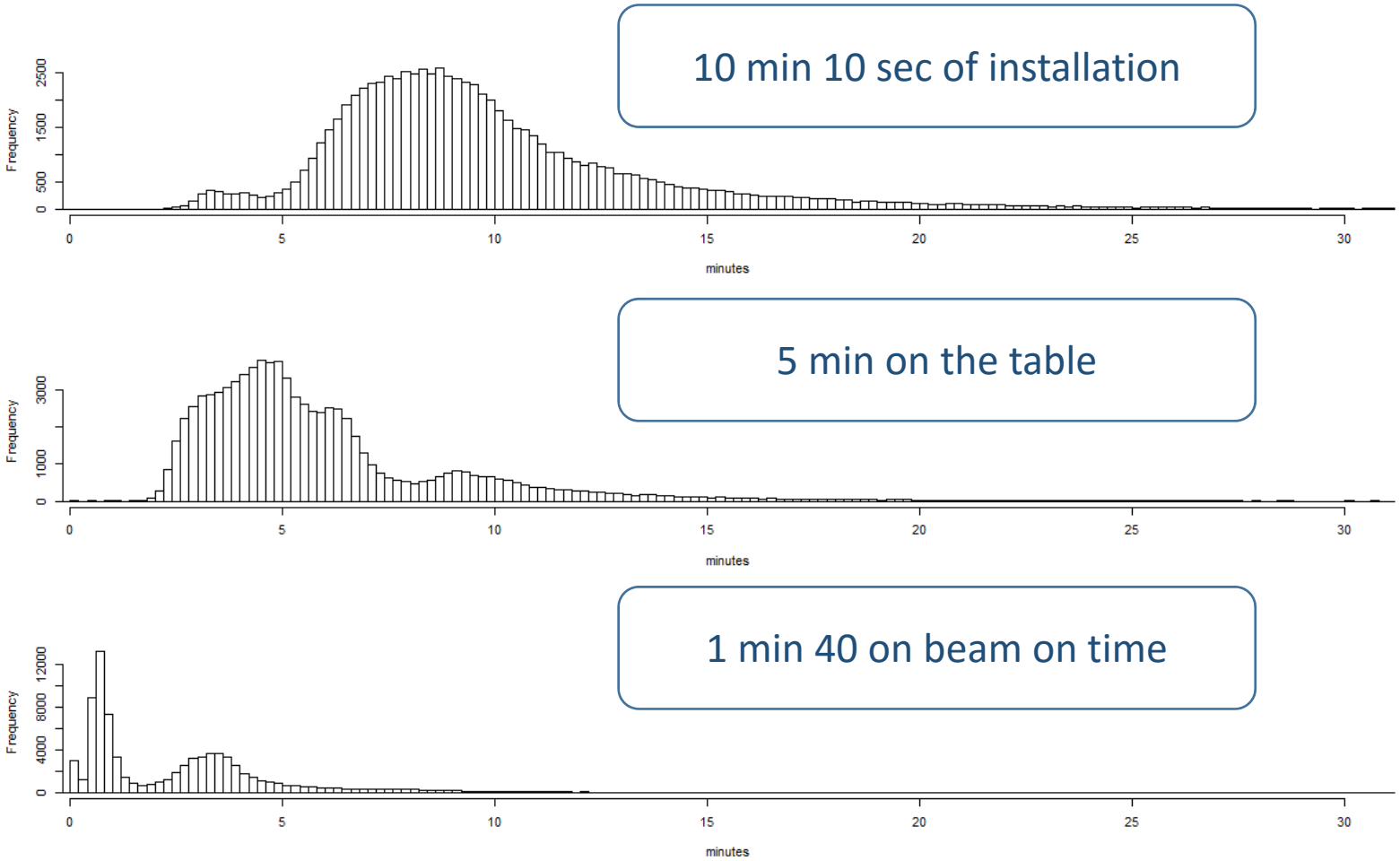


(Q3) Booking a LINAC for how long ?

- At the moment all patients are booked for 20 minutes on the LINACS (so schedules look nice).
- How long do each treatment really lasts ?
- Data driven approach (thanks to RFID)

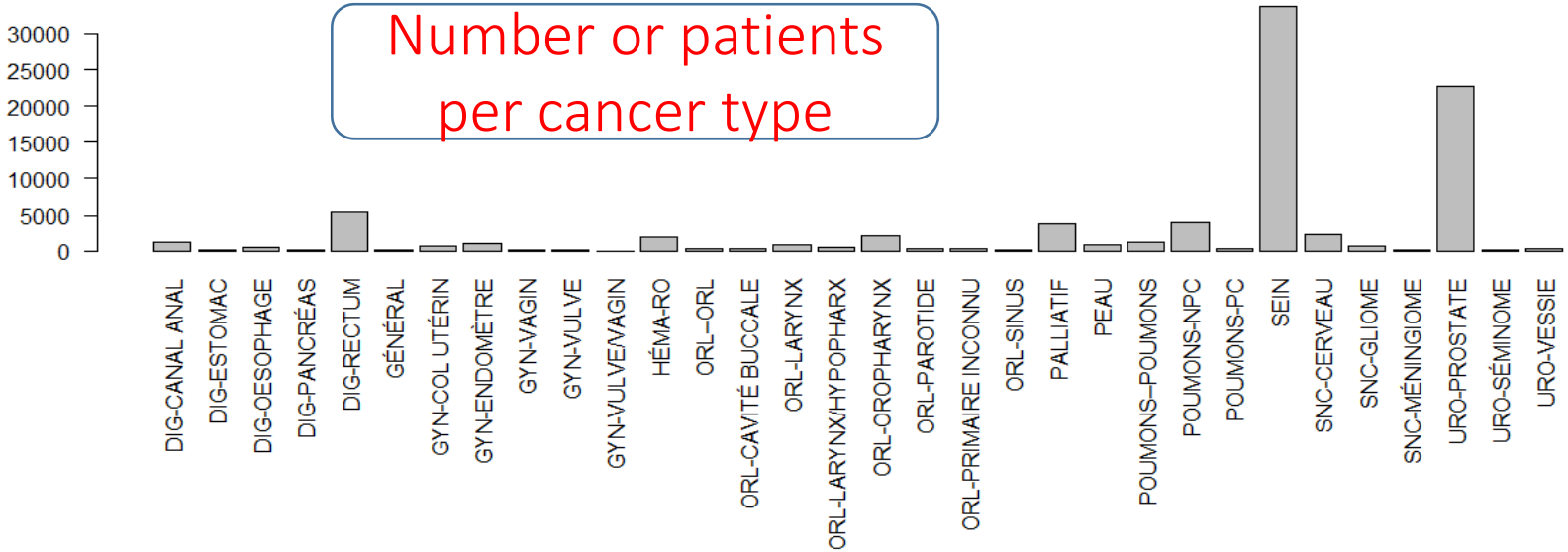
	A	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Dossier	Site prescrite	Activité	Lieu RV	Statut RV	Date RV	Heure RV	Heure d'arrivée du patient	Heure de vérification de la mise en place du patient en salle	Heure prise d'image CT	Heure traitée 1er champ	Heure de complétion RV	Lieu traité	Lieu traité diffère du lieu RV ?
2	1	P1SEIN D	TROEXTC	RO Salle 5	DC	18/03/2013	12:50:00	12:09:04	12:41:41	12:45:46	12:47:01	12:49:36	RO Salle 5	
3	1	P1SEIN D	TROEXTC	RO Salle 5	C	19/03/2013	12:50:00	12:20:44	12:55:49	12:59:16	13:00:34	13:03:16	RO Salle 5	
4	1	P1SEIN D	TROEXTC	RO Salle 5	C	20/03/2013	12:50:00	11:59:18	12:09:15	12:12:26	12:15:00	12:16:04	RO Salle 5	
5	1	P1SEIN D	TROEXTC	RO Salle 5	C	21/03/2013	12:50:00	11:47:14	12:24:38	00:00:00	12:34:26	12:35:04	RO Salle 5	
6	1	P1SEIN D	TROEXTC	RO Salle 5	C	22/03/2013	12:50:00	12:00:48	12:43:52	12:53:39	12:56:23	12:59:10	RO Salle 5	
7	1	P1SEIN D	TROEXTC	RO Salle 5	C	25/03/2013	12:50:00	12:02:30	12:44:32	12:48:25	12:51:07	12:51:43	RO Salle 5	
8	1	P1SEIN D	TROEXTC	RO Salle 5	C	26/03/2013	12:50:00	12:11:21	12:41:25	12:48:57	12:51:35	12:52:06	RO Salle 5	
9	1	P1SEIN D	TROEXTC	RO Salle 5	C	27/03/2013	12:50:00	12:11:31	12:42:03	12:44:47	12:47:21	12:48:02	RO Salle 5	
10	1	P1SEIN D	TROEXTC	RO Salle 5	C	28/03/2013	12:50:00	12:14:28	12:26:36	12:30:12	12:32:50	12:33:46	RO Salle 5	
11	1	P1SEIN D	TROEXTC	RO Salle 5	C	02/04/2013	12:50:00	11:51:22	12:11:20	12:15:44	12:18:21	12:18:49	RO Salle 5	
12	1	P1SEIN D	TROEXTC	RO Salle 5	C	03/04/2013	12:50:00	12:03:39	12:32:08	12:35:51	12:38:30	12:39:10	RO Salle 5	
13	1	P1SEIN D	TROEXTC	RO Salle 5	C	04/04/2013	12:50:00	11:55:03	12:39:21	12:43:22	12:45:58	12:46:37	RO Salle 5	
14	1	P1SEIN D	TROEXTC	RO Salle 5	C	05/04/2013	12:50:00	11:46:52	12:34:13	12:38:26	12:41:00	12:41:28	RO Salle 5	
15	1	P1SEIN D	TROEXTC	RO Salle 5	C	08/04/2013	12:50:00	11:56:38	12:23:03	12:30:19	12:32:57	12:33:31	RO Salle 5	
16	1	P1SEIN D	TROEXTC	RO Salle 5	C	09/04/2013	12:50:00	11:50:47	12:55:46	12:59:07	13:01:47	13:02:26	RO Salle 5	
17	1	P1SEIN D	TROEXTC	RO Salle 5	C	10/04/2013	12:50:00	11:51:47	12:57:33	13:01:45	13:04:20	13:05:09	RO Salle 5	
18	1	P2SEIN D (SURDO)	TROEXTC	RO Salle 5	C, PS	11/04/2013	12:50:00	11:56:17	13:26:35	00:00:00	13:28:09	13:28:34	RO Salle 5	
19	1	P2SEIN D (SURDO)	TROEXTC	RO Salle 5	C	12/04/2013	12:50:00	11:50:54	13:14:27	00:00:00	13:20:41	13:20:50	RO Salle 5	
20	1	P2SEIN D (SURDO)	TROEXTC	RO Salle 6	C	15/04/2013	16:30:00	16:11:49	16:27:44	00:00:00	16:35:35	16:35:43	RO Salle 6	
21	1	P2SEIN D (SURDO)	TROEXTC	RO Salle 5	FC	16/04/2013	12:50:00	11:47:14	12:45:48	00:00:00	12:51:16	12:51:25	RO Salle 5	
22	2	P1SUS-CLAV D	TROEXTC	RO Salle 2	DC	27/08/2013	11:50:00	11:35:24	12:04:21	12:10:14	12:11:43	12:21:43	RO Salle 2	
23	2	P1SUS-CLAV D	TROEXTC	RO Salle 2	C	28/08/2013	11:50:00	11:37:32	11:55:13	11:59:46	12:01:24	12:09:39	RO Salle 2	
24	2	P1SUS-CLAV D	TROEXTC	RO Salle 2	C	29/08/2013	11:50:00	11:02:11	12:12:19	12:17:50	12:18:50	12:25:47	RO Salle 2	
25	2	P1SUS-CLAV D	TROEXTC	RO Salle 2	C	30/08/2013	11:50:00	11:37:44	11:52:23	11:56:12	11:57:39	12:05:07	RO Salle 2	
26	2	P1SUS-CLAV D	TROEXTC	RO Salle 2	C	03/09/2013	11:50:00	11:38:06	12:07:06	12:11:21	12:12:24	12:20:12	RO Salle 2	
27	2	P1SUS-CLAV D	TROEXTC	RO Salle 2	C	04/09/2013	11:50:00	11:38:38	12:24:41	12:28:32	12:29:50	12:37:20	RO Salle 2	
28	2	P1SUS-CLAV D	TROEXTC	RO Salle 2	C	05/09/2013	11:50:00	11:36:04	12:13:07	12:17:05	12:18:10	12:26:06	RO Salle 2	

Duration profile



Looking at cancer type

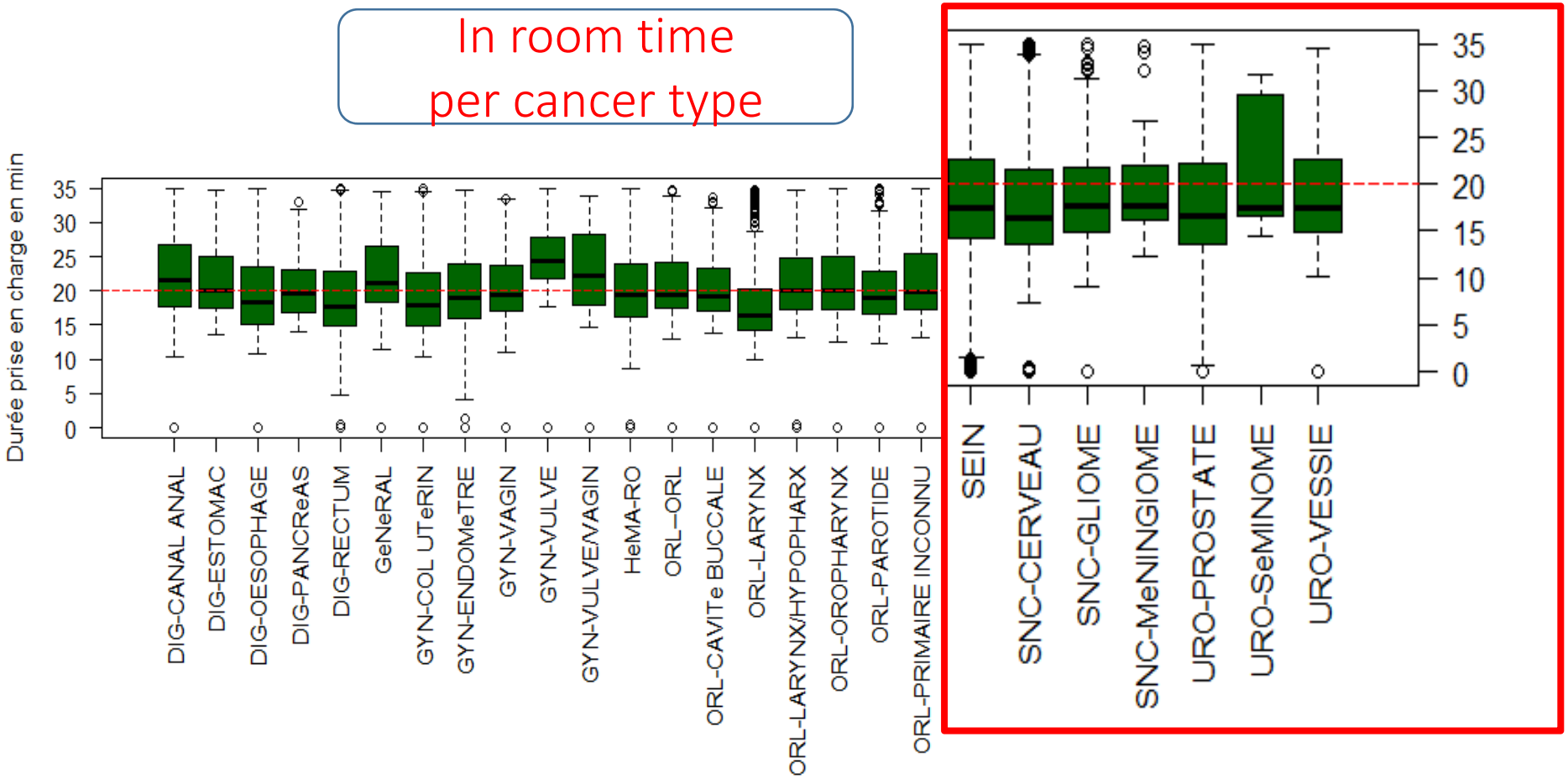
Going beyond the average…



Looking at cancer type

Going beyond the average…

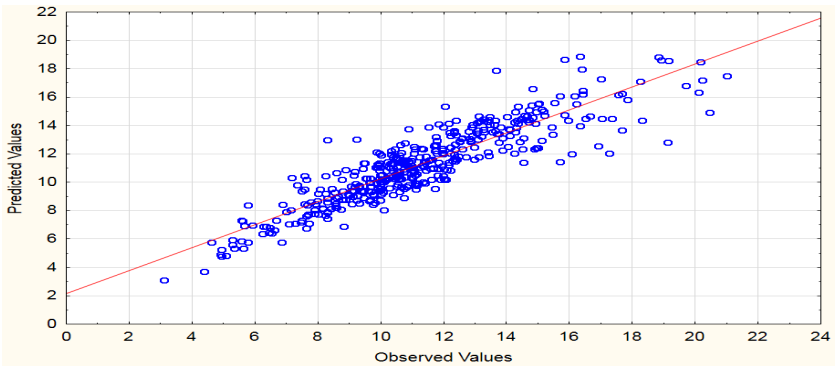
In room time
per cancer type



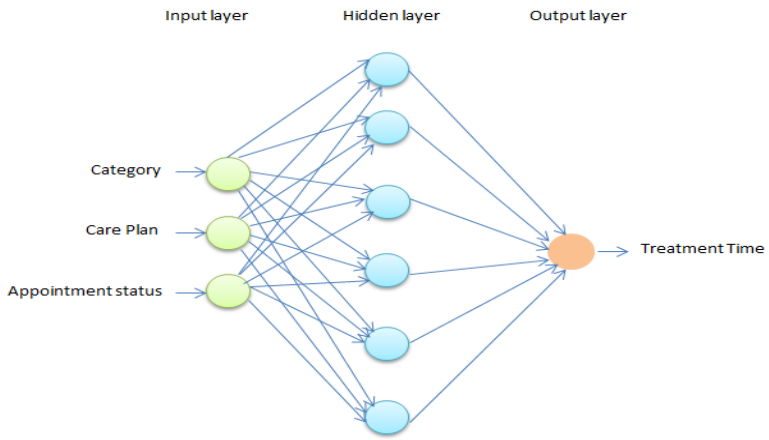
What to forecast ?

- We found that predicting a precise duration was difficult and not "actionable":
  - the hospital will not create time slots of 12m41 seconds…
- We instead try to predict whether a treatment should be given a time slot of a multiple of 5 minutes (5,10,15, ….)
  - There is a wish that these be aligned visually.

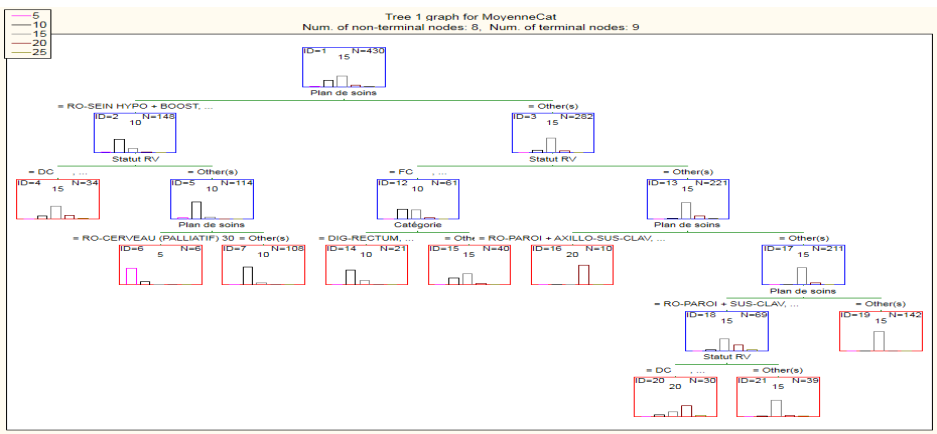
What works best ?



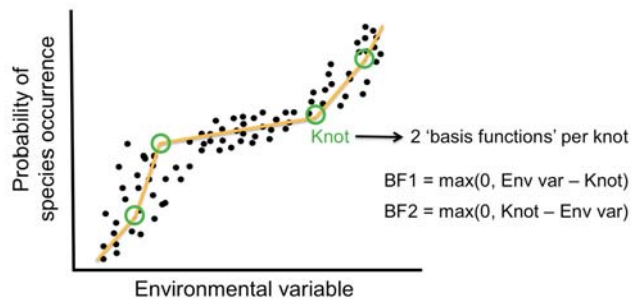
General linear model



Neural Network

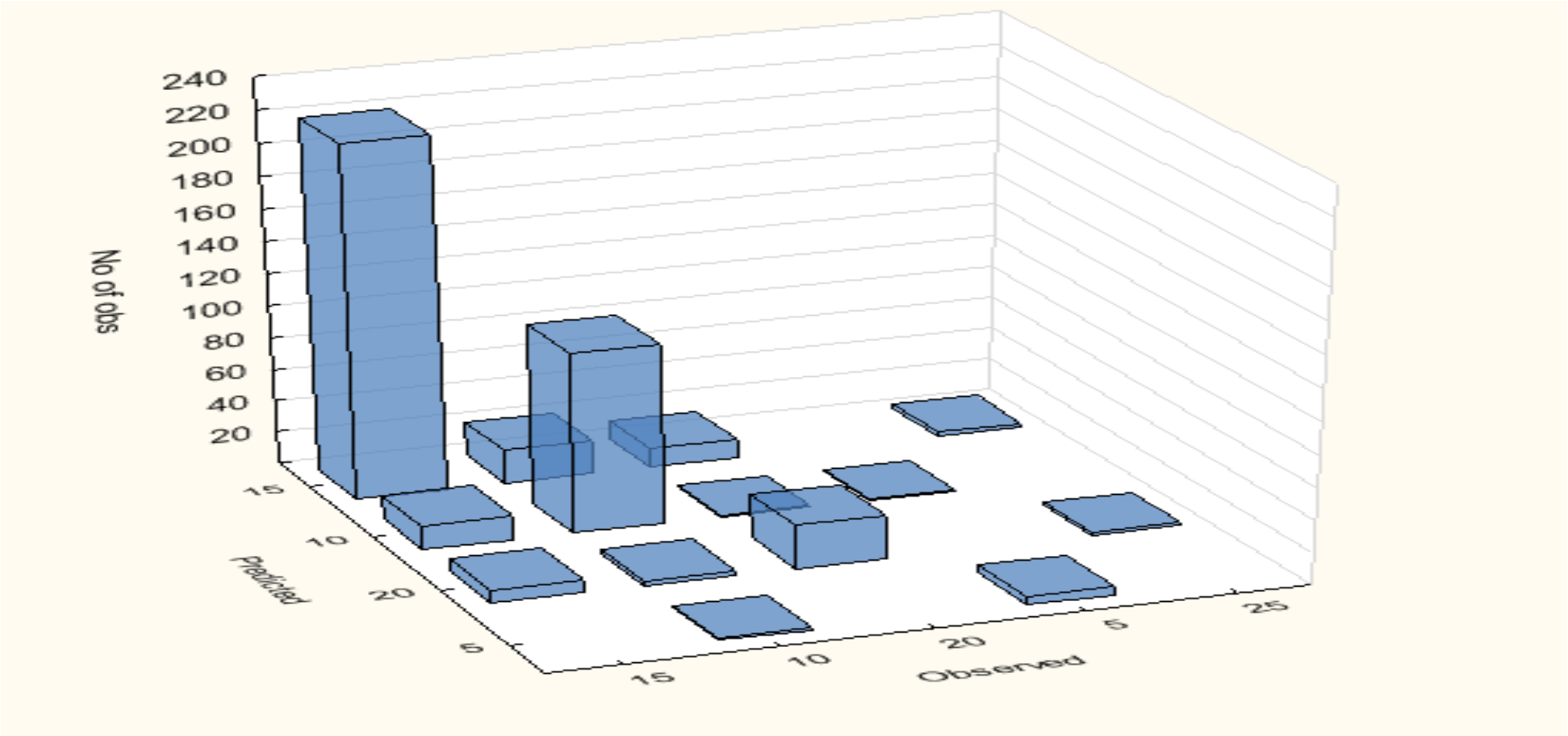


Classification Tree



Multivariate Adaptative Regression Splines(MARS)

Classification Tree



accuracy of **84%**

More precise treatment slots

- Can we build a calendar with different time slots for different cancer type ?
- Does cancer type really explains duration ?
- How to manage variance and uncertainty ?
- Can this be done without increasing the patient wait time (actual is 7 minutes) ?
- Data seems to indicate huge potential gain
- Solutions will to be tested through simulation

Still an ongoing project…

Conclusion

- Efficiency is key in fighting cancer given the continuously increasing
  - number of patients
  - cost of new technology

- Predictive (ML and stat) and Prescriptive (OR) Analytics can help to fully utilised those infrastructure

FOLLUS US @ HANALOG.CA

Warm Thanks to

Radio-oncologists and Physicists

MA Fortin - M Hinsse - B Carroza - D Roberge - S Bedaoui - D Craft – T Long

Students

M Taobane - A Legrain - M Mahnam – D Bentayeb

Colleagues

N Lahichi -- M Gendreau

Past and Present Partners

CIRRELT

Centre de santé et de services sociaux de Laval

Centre intégré de cancérologie

ELEKTA

CHUM