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

Louis-Martin Rousseau
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?

Local  |  Locally advanced  |  Metastatic

Spread of disease

Surgery

Radiotherapy

Chemotherapy

About 50% of cancer patients will receive radiotherapy
Tools

Internal

External

Vickers 6 Prototype Newcastle-on-Tyne 1960
Teams

<table>
<thead>
<tr>
<th></th>
<th>Chemotherapy</th>
<th>Radiotherapy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prescribes</td>
<td>Oncologist</td>
<td>Radiation Oncologist</td>
</tr>
<tr>
<td>Prepares</td>
<td>Pharmacist</td>
<td>Physicist</td>
</tr>
<tr>
<td>Delivers</td>
<td>Nurse</td>
<td>Therapist</td>
</tr>
</tbody>
</table>

Care Trajectory

Referral → NP appointment → CT SIM → Planning → Treatment → FUP
Care Trajectory in details
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

<table>
<thead>
<tr>
<th></th>
<th>Palliative</th>
<th>Curative 1</th>
<th>Curative 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time Interval</td>
<td>&lt; 3 days</td>
<td>&lt; 14 days</td>
<td>&lt; 28 days</td>
</tr>
</tbody>
</table>
Different possible approaches

Stochastic Optimization

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

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

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.
Structure of the model

\[ r_j = 7 \quad d_j = \text{Maximum delay} = 14 \]

Pattern 0

\[ c_{0j} = 8 \times 1 \]

Pattern 1

\[ c_{1j} = 16 \times 1 + 2 \times 50 \]

Treatment planning fix to 7 days ... for now
Stochastic Programming Model

\[
\min \sum_{i \in S_j} c_{ij} x_{ij} + E_{\omega \in \Omega_j} \left[ \sum_{l \in P^\omega} \sum_{i \in S_l} c_{il} y_{il}^\omega + \sum_{k \in H} \sum_{m \in M} c^o z_{mk}^\omega \right]
\]

subject to:

\[
\sum_{i \in S_j} x_{ij} = 1
\]

\[
\sum_{i \in S_l} y_{il}^\omega = 1, \quad \forall \omega \in \Omega_j, \forall l \in P^\omega
\]

\[
\sum_{i \in S_j} a_{ijk} x_{ij} + \sum_{l \in P^\omega} \sum_{i \in S_l} a_{ikl} y_{il}^\omega \leq F_k^m + z_{mk}^\omega, \quad \forall m \in M, \forall k \in H, \forall \omega \in \Omega_j
\]

\[
\mathbb{1}_{P_p(j)} \sum_{i \in S_j} a_{ijk} x_{ij} + \sum_{l \in P^\omega} \sum_{i \in S_l} a_{ikl} y_{il}^\omega \geq z_{mk}^\omega, \quad \forall m \in M, \forall k \in H, \forall \omega \in \Omega_j
\]

\[
\sum_{k=b}^{b+4} z_{mk}^\omega \leq O_{\text{week}}, \quad \forall m \in M, \forall b \in B, \forall \omega \in \Omega_j
\]

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

\[
x_{ij} \in \{0, 1\}, \quad \forall i \in S_j
\]

\[
y_{il}^\omega \in \{0, 1\}, \quad \forall l \in P^\omega, \forall i \in S_l, \forall \omega \in \Omega_j
\]

Choose greedily the pattern
With best reduced cost

Pattern 1
Pattern 0
Initial Results

<table>
<thead>
<tr>
<th>Due date violations</th>
<th>Average waiting time</th>
<th>Utilization</th>
<th>Overtime</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;3</td>
<td>&gt;14</td>
<td>&gt;28</td>
<td></td>
</tr>
<tr>
<td>CICL</td>
<td>14</td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td>OSCO - 1</td>
<td>9</td>
<td>6</td>
<td>0</td>
</tr>
</tbody>
</table>

CICL real data:
- 170 patients;
- 120 days;
- 2 linacs with 23 slots.

(Q2) How much for treatment preparation?

Appointment booking  Unknown dosimetry duration  Preparation completed
Preparation Tasks

Contouring
- Contour OAR
- Contour Physician

Planning
- Plan (Coordinator)
- Plan check (Check location)
- Plan approval
- Review Plan (Dosimetrist A)
- Re Plan (Dosimetrist A)
- Plan approval

Validation
- Export Plan to MOSAIQ
- Verify MU (Physics)
- Chart Prep (Linacs)
- Chart check
- Final check
- Ready to treat

All of which is happening in absence of the patient
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

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

1. Start
2. New patient?
   - YES: Find earliest feasible appointment
     - Type of patient?
       - PALLIATIVE: Choose earliest feasible appointment
       - CURATIVE: Generate patient scenarios
         - Initialize $x_{ij} = 0$
         - Solve the relaxed subproblems
         - Generate dosimetry planning columns
   - NO: Increment day
3. Exceed horizon?
   - YES: End
   - NO: Get linacs and dosimetrists state from IS
        - Find earliest feasible appointment
        - Solve the restricted stochastic matching problem
          - Update appointment costs
            - Add earliest starting time
        - Relay the appointment to IS
        - Add new columns?
          - YES: End
          - NO: Update appointment costs
            - Add earliest starting time

Information System (IS)
New Results

<table>
<thead>
<tr>
<th>Cancellations</th>
<th>Due date (in days)</th>
<th>Average waiting time</th>
<th>Overtime</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&gt;3</td>
<td>&gt;14</td>
<td>&gt;28</td>
</tr>
<tr>
<td>CICL</td>
<td>230</td>
<td>373</td>
<td>104</td>
</tr>
<tr>
<td>OSCO 1</td>
<td>107</td>
<td>335</td>
<td>67</td>
</tr>
<tr>
<td>OSCO 2</td>
<td>1</td>
<td>326</td>
<td>119</td>
</tr>
</tbody>
</table>

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.*
(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)
Duration profile

10 min 10 sec of installation

5 min on the table

1 min 40 on beam on time
Looking at cancer type

Going beyond the average...

Number or patients per cancer type
Looking at cancer type

Going beyond the average...

In room time per cancer type

Durée prise en charge en min

0 5 10 15 20 25 30 35


SEIN SNC-CERVEAU SNC-GLIOME SNC-MENINGE URO-PROSTATE URO-SEMINOME URO-VESIQUE

Figure 26: Durée de prise en charge des patients en fonction de leur catégorie

Dans la mesure où les catégories SEIN et URO-PROSTATE représentent à eux seuls 66% des patients et qu'il semblerait que la durée de prise en charge soit inférieure au 20 minutes théoriques, il est intéressant de voir si tous les patients sont homogènes, ce qui ne semble pas être le cas au vu de la dispersion pour ces catégories, et si des critères autres peuvent être identifiés.

Concernant le cancer du sein, si l'on se concentre sur les plans de soin représentant 90% des traitements, 3 catégories de patients peuvent être répertoriées: ceux dont la prise en charge est inférieure généralement à 20 minutes (RO SEIN HYPO, RO SEIN HYPO + BOOST et RO SEIN) et qui représentent 54% des patients, ceux dont la prise en charge est d'environ 20 minutes (RO SEIN + BOOST et RO SEIN + AXILLO SUS CLAV + BOOST) et qui représentent 24% des patients et finalement ceux dont la prise en charge est supérieure à 20 minutes (RO PAROI + AXILLO SUS CLAV et RO PAROI + SUS CLAV) et qui représentent 12% des patients. Cependant, on constate que l'écart type est assez élevé; ainsi, d'autres critères doivent rentrer en jeu dans l'explication de la durée de traitement.
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...
(Q4) 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 $(5 \times 10)$ beam and 100 sectors,
  - there are $7.1 \times 10^{251}$ apertures shapes.
- Column generation (CG) is a leading optimization technique successful solving large-scale problems
  - Exploits decomposable structures
  - Handles large number of variables

Restricted Master Model

Model Information

New promising columns

Pricing Problem
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 described 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

\[ \text{GP : } \min \quad \mathbf{F}(z) + w T_{\text{max}} \quad \text{Weighted quality and time objective} \]

\[ z_j = \sum_{k \in K} \sum_{h \in H_k} D_{j,h}(A_{h}^k) y^k \rho_h t_h \quad \forall j \in \mathcal{V} \]

\[ \sum_{k \in K} a_h^k y^k \leq 1 \quad \forall h \in H \]

\[ |\rho_{h+1} - \rho_h| \leq \Delta_\rho \quad \forall h = 1, 2, \cdots, |H| - 1 \]

\[ 0 \leq \rho_h \leq R \quad \forall h \in H \]

\[ \sum_{k \in K} \tau_{h,h+1}^k y^k \leq t_h \quad \forall h \in H \]

\[ |t_{h+1} - t_h| \leq \Delta_t \quad \forall h = 1, 2, \cdots, |H| - 1 \]

\[ T \leq t_h \leq \mathcal{T} \quad \forall h \in H \]

\[ \sum_{h \in H} t_h \leq T_{\text{max}} \]

\[ y^k \in \{0, 1\} \quad \forall k \in K \]
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 ≤ 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 represented by a vector in space $\mathbb{R}^n$ ($n =$ number of beamlets).

![Diagram of steps](https://via.placeholder.com/150)
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.

<table>
<thead>
<tr>
<th>Iter</th>
<th># Transfer</th>
<th>Iter Time (Sec.)</th>
<th>Avg.Dist</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>54.27054</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>19265</td>
<td>0.952947</td>
<td>11.57258</td>
</tr>
<tr>
<td>2</td>
<td>2160</td>
<td>0.74742</td>
<td>10.51782</td>
</tr>
<tr>
<td>3</td>
<td>301</td>
<td>0.733188</td>
<td>10.44499</td>
</tr>
<tr>
<td>4</td>
<td>40</td>
<td>0.729702</td>
<td>10.44053</td>
</tr>
<tr>
<td>5</td>
<td>13</td>
<td>0.738105</td>
<td>10.43969</td>
</tr>
</tbody>
</table>
Experimental evaluation

<table>
<thead>
<tr>
<th>Case Characteristics</th>
<th>Algorithm Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total # beamlets</td>
<td>Max dose rate 600 MU/min</td>
</tr>
<tr>
<td>Beamlet size (mm)</td>
<td>Max leaf speed 3 cm/sec</td>
</tr>
<tr>
<td>Voxel resolution (mm)</td>
<td>Max fluence change 2 MU/s</td>
</tr>
<tr>
<td># Target voxels</td>
<td>Max time change 2 s</td>
</tr>
<tr>
<td># Body voxels</td>
<td>Gantry speed [1 6]°/sec</td>
</tr>
</tbody>
</table>

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

Results

The computational time of Plan 100% - 100% and Plan 10%-5% are about 22.25 and 3.75 min, respectively.

Full plan 100%-100% vs. aggregated plan 10%-5%

(CUSM, November 2016)

VMAT treatment planning

CPU Times
Agg: 3m75s
Full: 22m25s
Effect of delivery time

Figure 3: DVH plots for Plan-Opt (2.61 minutes), Plan-2 (2 minutes), and Plan-6 (6 minutes) obtained for the prostate case using GCRGH. (a) Comparison of Plan-Opt and Plan-2, and (b) comparison of Plan-Opt and Plan-6. The solid lines indicate Plan-Opt; the dashed lines indicate Plan-2 and Plan-6.
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 utilise 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 - P Descroix - V Cohen - F Camiat

Colleagues
N Lahichi - P Baptiste - M Gendreau

Past and Present Partners