Adaptive and Robust Radiation Therapy Optimization for Lung Cancer

Velibor V. Mišić [†] Timothy C. Y. Chan [†]

[†] Department of Mechanical and Industrial Engineering, University of Toronto

November 15, 2011

INFORMS Annual Meeting 2011, Charlotte, NC

Lung cancer

- Leading cause of cancer death in the US and Canada
- Over 180,000 deaths this year; 25% of cancer deaths

Intensity modulated radiation therapy (IMRT)

- Very popular form of radiation therapy
- Several beams; each beam consists of smaller beams or beamlets
- Basic problem: find beamlet intensities that deliver at least some prescription dose to the tumour at minimal healthy tissue damage

Uncertainty

- For lung cancer, most significant uncertainty comes from breathing motion
- Patient breathes; tumour is not in the same position during treatment session
- Breathing pattern is not known precisely

Approaches to uncertainty

- **Nominal approach**: assume patient will breathe according to a single breathing pattern
  - If actual breathing pattern is different from planned, tumour underdose is very likely
- **Margin approach**: assume patient can breathe according to any breathing pattern
  - Tumour dose is guaranteed to be sufficient, but cost to healthy tissue is high
- **Robust optimization**: Given a set of breathing patterns, find the treatment that minimizes damage to healthy tissue and meets tumour dose constraints under those breathing patterns (Bortfeld et al. 2008)

Robust optimization - uncertainty

- X: a set of breathing motion states
- $\mathbf{p} = (p(x))_{x \in X}$: a breathing motion probability mass function (PMF)
- $p(x)$: proportion of time patient spends in state x during a treatment session
- P: uncertainty set; collection of $\mathbf{p}$ vectors that we wish to protect ourselves against

Robust optimization - decision variables and parameters

- $\mathcal{B}$: set of beamlets to be used for treatment
- w_b: intensity of beamlet $b \in \mathcal{B}$
- $\mathcal{V}$: set of all voxels; $\mathcal{T}$: set of tumour voxels
- For each voxel v, motion state x and beamlet b, a dose deposition coefficient $\Delta_{v,x,b}$
- Dose to voxel v under PMF $\mathbf{p}$:

$$\sum_{x \in X} \sum_{b \in \mathcal{B}} \Delta_{v,x,b} p(x) w_b$$

- θ_v: minimum prescription dose of tumour voxel $v \in \mathcal{T}$

Robust optimization - model

Formulation is

$$
\begin{aligned}
\text{minimize} \quad & \sum_{v\in\mathcal{V}} \sum_{x\in X} \sum_{b\in\mathcal{B}} \Delta_{v,x,b}\bar{p}(x) w_b \\
\text{subject to} \quad & \sum_{x\in X} \sum_{b\in\mathcal{B}} \Delta_{v,x,b} p(x) w_b \geq \theta_v, \quad \forall\ v \in \mathcal{T},\ \forall\ \mathbf{p} \in P, \\
& \sum_{x\in X} \sum_{b\in\mathcal{B}} \Delta_{v,x,b} p(x) w_b \leq \gamma\theta_v, \quad \forall\ v \in \mathcal{T},\ \forall\ \mathbf{p} \in P, \\
& w_b \geq 0, \quad \forall\ b \in \mathcal{B},
\end{aligned}
$$

where $\bar{\mathbf{p}}$ is a PMF representative of the patient's overall breathing and $\gamma \geq 1$.

Robust optimization - properties

- If patient's PMF $\mathbf{p}$ is in P while $\mathbf{w}^*$ (optimal solution for RO problem with P) is being delivered, then

$$d_v = \sum_{x \in X} \sum_{b \in \mathcal{B}} \Delta_{v,x,b} p(x) w_b^* \geq \theta_v$$

  for every tumour voxel v
- Generally, the larger P is, the more dose is delivered to healthy tissue under $\mathbf{w}^*$. (If $P^1 \supseteq P^2$, then

$$\text{Obj. value with } P^1 \geq \text{Obj. value with } P^2$$

  holds)

... so what's the problem?

Fractionation:

- Prescribed dose is divided up into n smaller amounts (fractions)
- Each fraction is delivered once a day every day for a period of 4 - 6 weeks
- Healthy tissue heals faster than cancerous tissue

Fractionation (continued)

- In Bortfeld et al. (2008), planner solves one problem before start of treatment, gets $\mathbf{w}$, delivers $\mathbf{w}/n$ in every fraction (static) ...
- ... but what if patient's $\mathbf{p}$ changes over treatment?
  - Patient may be nervous in the beginning, but become more relaxed by the end
  - Progression of disease may change patient's breathing
- ... what if uncertainty set P was determined inappropriately?
  - P may be very large, but patient's $\mathbf{p}$'s are actually tightly clustered around a single value

Adaptive robust optimization

Our idea: Solve a sequence of robust optimization problems (one for each fraction), with the uncertainty set updated using the most recent PMF each time, and deliver the resulting solutions

Our approach - I

Initialization:

1. Select an initial uncertainty set P^1
2. Solve the robust problem associated with P^1 to obtain beamlet weight vector $\mathbf{w}^1$ for the first fraction

Our approach - I

Initialization:

1. Select an initial uncertainty set P^1
2. Solve the robust problem associated with P^1 to obtain beamlet weight vector $\mathbf{w}^1$ for the first fraction

Our approach - II

In iteration (fraction) i ($\in \{1, \ldots, n\}$):

1. Deliver $\mathbf{w}^i/n$ to the patient
2. Measure the patient's breathing, obtain $\mathbf{p}^i$
3. Generate new uncertainty set P^{i+1} from previous set P^i and just observed $\mathbf{p}^i$
4. Solve the robust problem with P^{i+1} to obtain beamlet weight vector $\mathbf{w}^{i+1}$ for the next fraction
5. Set $i = i + 1$

Our approach - II

In iteration (fraction) i ($\in \{1, \ldots, n\}$):

1. Deliver $\mathbf{w}^i/n$ to the patient
2. Measure the patient's breathing, obtain $\mathbf{p}^i$
3. Generate new uncertainty set P^{i+1} from previous set P^i and just observed $\mathbf{p}^i$
4. Solve the robust problem with P^{i+1} to obtain beamlet weight vector $\mathbf{w}^{i+1}$ for the next fraction
5. Set $i = i + 1$

Our approach - II

In iteration (fraction) i ($\in \{1, \ldots, n\}$):

1. Deliver $\mathbf{w}^i/n$ to the patient
2. Measure the patient's breathing, obtain $\mathbf{p}^i$
3. Generate new uncertainty set P^{i+1} from previous set P^i and just observed $\mathbf{p}^i$
4. Solve the robust problem with P^{i+1} to obtain beamlet weight vector $\mathbf{w}^{i+1}$ for the next fraction
5. Set $i = i + 1$

Our approach - II

In iteration (fraction) i ($\in \{1, \ldots, n\}$):

1. Deliver $\mathbf{w}^i/n$ to the patient
2. Measure the patient's breathing, obtain $\mathbf{p}^i$
3. Generate new uncertainty set P^{i+1} from previous set P^i and just observed $\mathbf{p}^i$
4. Solve the robust problem with P^{i+1} to obtain beamlet weight vector $\mathbf{w}^{i+1}$ for the next fraction
5. Set $i = i + 1$

Our approach - II

In iteration (fraction) i ($\in \{1, \ldots, n\}$):

1. Deliver $\mathbf{w}^i/n$ to the patient
2. Measure the patient's breathing, obtain $\mathbf{p}^i$
3. Generate new uncertainty set P^{i+1} from previous set P^i and just observed $\mathbf{p}^i$
4. Solve the robust problem with P^{i+1} to obtain beamlet weight vector $\mathbf{w}^{i+1}$ for the next fraction
5. Set $i = i + 1$

Uncertainty sets

- The set of all probability distributions on X:

$$\mathcal{P} = \left\{ \mathbf{p} \in \mathbb{R}^{|X|} \;\middle|\; \forall\ x \in X,\ p(x) \geq 0;\ \sum_{x \in X} p(x) = 1 \right\}$$

- An uncertainty set P is specified by a lower bound vector $\boldsymbol{\ell}$ and upper bound vector $\mathbf{u}$:

$$P = \{\mathbf{p} \in \mathcal{P} \,|\, \forall\ x \in X,\ \ell(x) \leq p(x) \leq u(x)\}\,.$$

Updating the uncertainty set

- Exponential smoothing:

$$\boldsymbol{\ell}^{k+1} = (1-\alpha)\boldsymbol{\ell}^{k} + \alpha \mathbf{p}^{k}$$
$$\mathbf{u}^{k+1} = (1-\alpha)\mathbf{u}^{k} + \alpha \mathbf{p}^{k}$$

where $\alpha \in [0, 1]$.

- Running average:

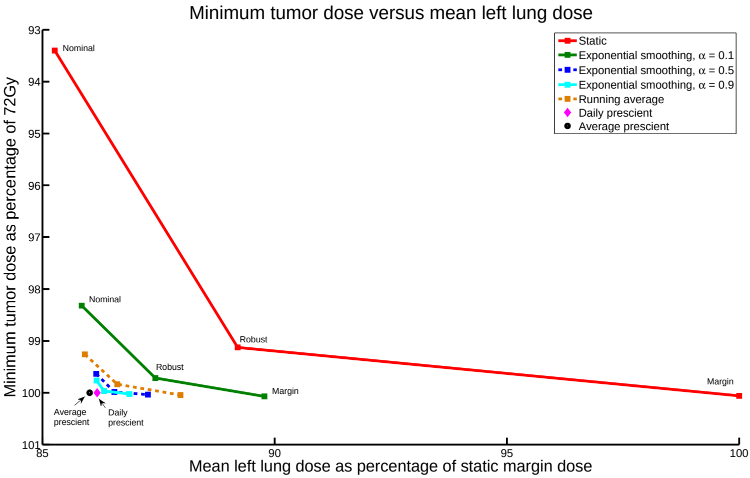
$$\boldsymbol{\ell}^{k+1} = \frac{1}{k+1}(\boldsymbol{\ell}^{1} + \sum_{i=1}^{k} \mathbf{p}^{i})$$

$$\mathbf{u}^{k+1} = \frac{1}{k+1}(\mathbf{u}^{1} + \sum_{i=1}^{k} \mathbf{p}^{i})$$

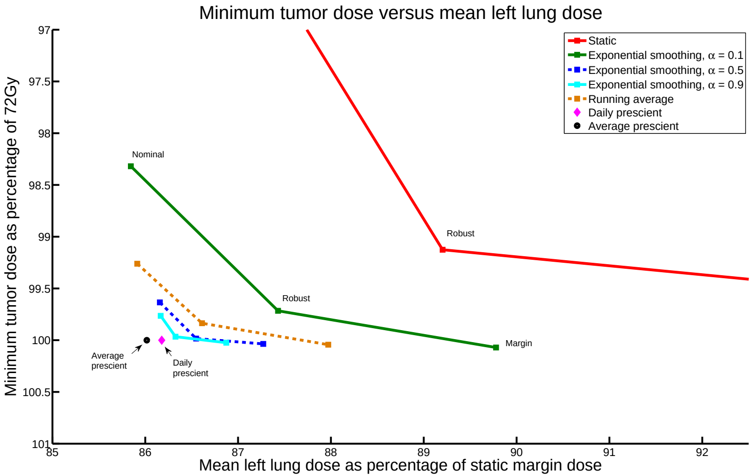
Computational results - I

- Sequence of real patient PMFs: $\mathbf{p}^1, \mathbf{p}^2, \ldots, \mathbf{p}^n$
- Select initial uncertainty set
  - Nominal; $P = \{\tilde{\mathbf{p}}\}$
  - Margin; $P = \mathcal{P}$
  - Robust; in between $\{\tilde{\mathbf{p}}\}$ and $\mathcal{P}$.

Computational results - first PMF sequence



Computational results - first PMF sequence (zoomed in)



Computational results - takeaways

- Adaptive solutions are in general better than static solutions
- Choice of initial uncertainty set is less important than it is for static robust method
- Solutions are very close in quality to the prescient solution

Theoretical results

Suppose

$$\mathbf{p}^n \to \mathbf{p}^*$$

as $n \to \infty$. What can we say about

$$\frac{1}{n} \sum_{i=1}^{n} \Delta \mathbf{p}^i \mathbf{w}^i$$

as $n \to \infty$?
$(\Delta \mathbf{pw} = [\sum_{x \in X} \sum_{b \in \mathcal{B}} \Delta_{v,x,b} p(x) w_b]_{v \in \mathcal{V}})$

Convex-convergent update algorithms

Call an uncertainty set update algorithm convex-convergent if:

1. Each update is a convex combination of the most recent $\mathbf{p}$ and most recent $\boldsymbol{\ell}/\mathbf{u}$: for every $n \in \mathbb{Z}_+$, there exists $\alpha_n \in [0,1]$ such that

$$\boldsymbol{\ell}^{n+1} = (1-\alpha_n)\boldsymbol{\ell}^n + \alpha_n \mathbf{p}^n,$$
$$\mathbf{u}^{n+1} = (1-\alpha_n)\mathbf{u}^n + \alpha_n \mathbf{p}^n.$$

2. The updates inherit the convergence of the PMF sequence:

$$\mathbf{p}^n \to \mathbf{p}^* \Rightarrow \boldsymbol{\ell}^n \to \mathbf{p}^*, \mathbf{u}^n \to \mathbf{p}^*.$$

Optimal solution sets

Let

- $\mathbf{w}^*(\mathbf{p}^*)$ be the set of optimal solutions to the robust problem with $P = \{\mathbf{p}^*\}$, and
- $\mathbf{w}^*(\ell, \mathbf{u})$ be the set of optimal solutions to the robust problem with P defined by ℓ and $\mathbf{u}$.

Optimal dose distribution set

- Let
$$\mathbf{D} = \{\mathbf{d} \mid \mathbf{d} = \Delta \mathbf{p}^* \mathbf{w} \text{ for some } \mathbf{w} \in \mathbf{w}^*(\mathbf{p}^*)\}.$$
- $\mathbf{D}$ is the set of dose distributions that are obtained when a $\mathbf{w}$ from $\mathbf{w}^*(\mathbf{p}^*)$ is delivered and $\mathbf{p}^*$ is realized during delivery
- Every $\mathbf{d} \in \mathbf{D}$ meets the minimum tumour dose constraint (for every $v \in \mathcal{T}$, $d_v \geq \theta_v$)

Convergence of dose distributions

Theorem

Suppose $\mathbf{p}^n \to \mathbf{p}^*$ *and* $(\boldsymbol{\ell}^n)_{n=1}^{\infty}$ *and* $(\mathbf{u}^n)_{n=1}^{\infty}$ *are obtained by a convex-convergent update algorithm. Suppose* $\mathbf{w}^i \in \mathbf{w}^*(\boldsymbol{\ell}^i, \mathbf{u}^i)$ *for each* i*.*

Then for every $\epsilon > 0$*, there exists an* $N \in \mathbb{Z}_+$ *such that for all* $n > N$*,*

$$\frac{1}{n}\sum_{i=1}^{n} \Delta \mathbf{p}^i \mathbf{w}^i \in U(\mathbf{D}, \epsilon),$$

where

$$U(A, \epsilon) = \bigcup_{x \in A} B(x, \epsilon),$$

and $B(x, \epsilon)$ *is the* ϵ*-ball at* x*.*

Convergence of dose distributions (continued)

- For daily and average prescient, same theorem holds
- For static robust:

$$\frac{1}{n}\sum_{i=1}^{n} \Delta \mathbf{p}^{i}\mathbf{w} = \Delta \sum_{i=1}^{n} \frac{\mathbf{p}^{i}}{n}\mathbf{w} \rightarrow \Delta \mathbf{p}^{*}\mathbf{w}$$

as $n \rightarrow \infty$.

- $\Delta \mathbf{p}^{*}\mathbf{w}$ may have some underdosed tumour voxels, depending on where $\mathbf{p}^{*}$ is with respect to the uncertainty set P of $\mathbf{w}$

Summary

- Adaptive robust method greatly improves on the static robust method
- Method achieves performance comparable to optimal prescient algorithms
- Simple and does not require a large amount of information pre-treatment (but does require work during treatment)

Future work

- Reducing the frequency of adaptation
- Adaptation in a distributionally-robust setting

Acknowledgements

- NSERC and CIHR for financial support
- Dr. Thomas Bortfeld at MGH for patient data

Thank you for listening!

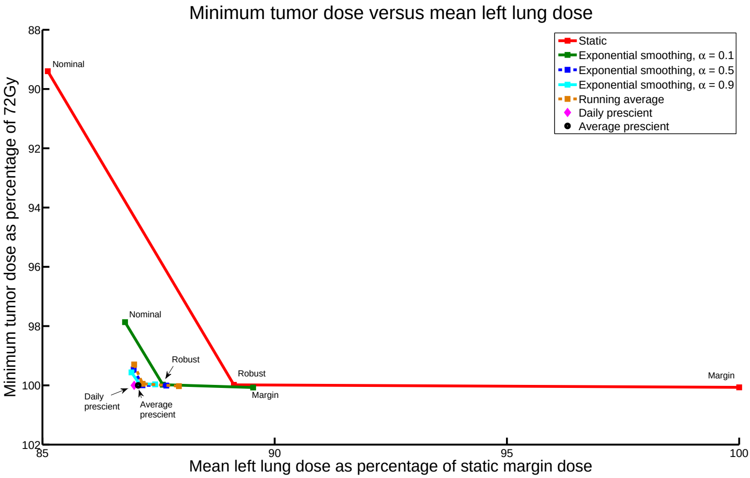
- Questions/comments?

Extras

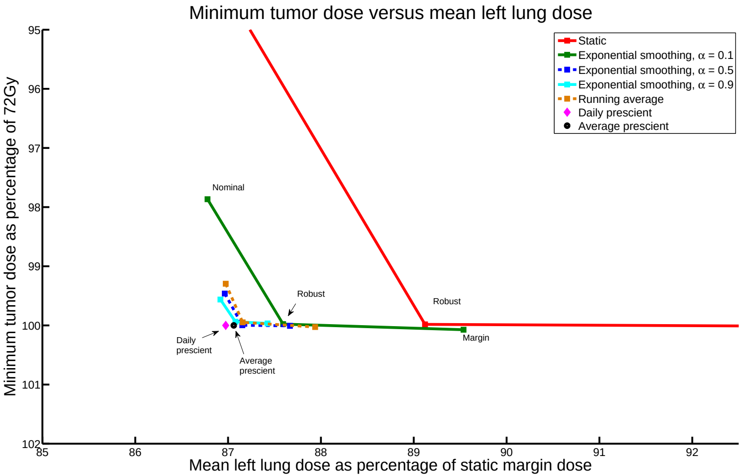
Prescient algorithms

- To analyze our results later, we also consider "prescient" algorithms
- Daily prescient algorithm:
  - On day i, set $\boldsymbol{\ell}^i = \mathbf{u}^i = \mathbf{p}^i$, so $P^i = \{\mathbf{p}^i\}$
  - On each day, tumour voxel v receives at least θ_v/n, so by end it receives at least θ_v
- Average prescient algorithm:
  - Calculate the average PMF: $\mathbf{p}_{\text{avg}} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{p}^i$.
  - On day i, set $\boldsymbol{\ell}^i = \mathbf{u}^i = \mathbf{p}_{\text{avg}}$, so $P^i = \{\mathbf{p}_{\text{avg}}\}$
  - By end of treatment, each tumour voxel v receives at least θ_v

Computational results - second PMF sequence



Computational results - second PMF sequence (zoomed in)



Reactive method

Define $\mathbf{d}^1$ as $d_v^1 = \theta_v$ for $v \in \mathcal{T}$, $\mathbf{u}^1$ as $u_v^1 = \gamma\theta_v$ for $v \in \mathcal{T}$. Start with $i = 1$.

1. Solve to obtain $\mathbf{w}^i$:

$$\begin{aligned}
\text{minimize} \quad & \sum_{v\in\mathcal{V}} \sum_{x\in X} \sum_{b\in\mathcal{B}} \Delta_{v,x,b}\bar{p}(x) w_b \\
\text{subject to} \quad & \sum_{x\in X} \sum_{b\in\mathcal{B}} \Delta_{v,x,b} p(x) w_b \geq d_v^i, \quad \forall v \in \mathcal{T},\ \mathbf{p} \in P, \\
& \sum_{x\in X} \sum_{b\in\mathcal{B}} \Delta_{v,x,b} p(x) w_b \leq u_v^i, \quad \forall v \in \mathcal{T},\ \mathbf{p} \in P, \\
& w_b \geq 0, \quad \forall b \in \mathcal{B}.
\end{aligned}$$

2. Deliver $\mathbf{w}^i/(n - i + 1)$.
3. Observe $\mathbf{p}^i$.

Reactive method (continued)

4. Set $\mathbf{d}^{i+1}$ as

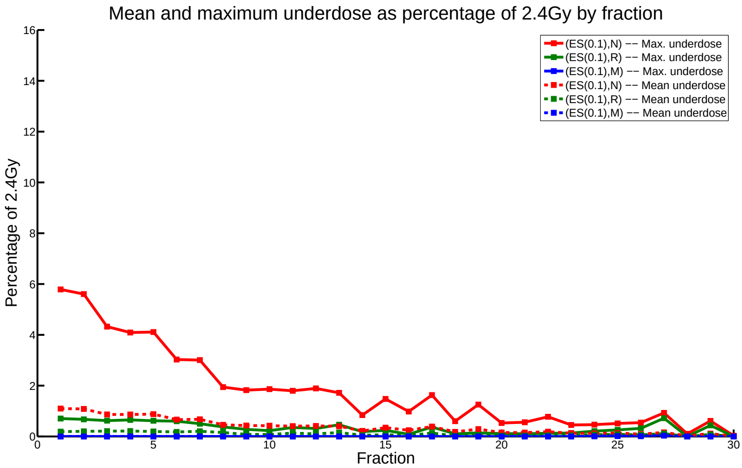
$$d_v^{i+1} = \max\{0, d_v^i - \sum_{x \in X} \sum_{b \in \mathcal{B}} \Delta_{v,x,b} p^i(x) w_b^i / (n - i + 1)\},$$

$\mathbf{u}_v^{i+1}$ as

$$u_v^{i+1} = \max\{0, u_v^i - \sum_{x \in X} \sum_{b \in \mathcal{B}} \Delta_{v,x,b} p^i(x) w_b^i / (n - i + 1)\}.$$

5. Generate P^{i+1} from P^i and $\mathbf{p}^i$.
6. Set $i = i + 1$.

Mean and max. underdose by fraction – non-reactive



Mean and max. underdose by fraction – reactive