Does software make me more efficient?

Mark J. Gooding, Ph.D.
Chief scientist
Mirada Medical
Presentation structure

• A bit about me

• A bit about the topic of discussion

• A look at efficiency in three clinical tasks
  – Auto-contouring
  – Multi-modal target contouring
  – Adaptive therapy

Very High Energy Electrons (>100 Mev) in Radiation Therapy
Colleen DesRosiers, PhD, DABR
But first…. Let’s get interactive

http://aamd2017.participoll.com
But first.... Let’s get interactive

What shape is a football?

http://aamd2017.participoll.com
But first.... Let’s get interactive

What shape is a football?

A

http://aamd2017.participoll.com

B
Why should you bother listening to me?

- 9 years in academic medical imaging research
- 8 years in commercial medical imaging research
- Chief Scientist

No experience as a dosimetrist
Expertise in medical software development
Passionate about what I/we do

Royal United Hospital Bath NHS Trust
Introduction to the question

*Does software make me more efficient?*
Introduction to the question

*Does software make me more efficient?*

How do you feel about your workflow?

A. Software is key to our efficient workflow

B. Our workflow is efficient despite the software

C. Workflow could be improved, but the software is not the main problem

D. Software hinders our workflow most

http://aamd2017.participoll.com
Introduction to the question

Does software make me more efficient?

Technology Adoption Curve

EVERETT ROGERS - DIFFUSION OF INNOVATIONS 1962

- 2.5% Innovators
- 13.5% Early Adopters
- 34% Early Majority Adopters
- 34% Late Majority Adopters
- 16% Laggards
Introduction to the question

*Does software make me more efficient?*

Over enthusiastic marketing
Auto-contouring
Atlas-based auto-contouring
Atlas-based auto-contouring

Deformable registration

Atlas (contoured CT) -- Current case
Atlas based auto-contouring

Single-atlas auto-contouring

DIR and contour warping

Contour consensus

Patient image

Multi-atlas auto-contouring

DIR, contour warping

Patient image

Selection of single atlas

DIR and contour warping

Contour consensus

Patient image

Selection of multiple atlases

DIR, contour warping

Patient image
Where are we on the adoption curve?

Does your institution **have** an auto-contouring system?

- A Yes
- B No
- C Don’t know

Where are we on the adoption curve?

Auto-contouring

- Reasonably good workflow integration
- Performance standardising
- Vendors competing on price / workflow integration

2.5% INNOVATORS
13.5% EARLY ADOPTERS
34% EARLY MAJORITY ADOPTERS
34% LATE MAJORITY ADOPTERS
16% LAGGARDS
Does it do it’s job?

Does your institution **routinely use** an auto-contouring system?

- A  Yes
- B  No
- C  Don’t know

http://aamd2017.participoll.com
Disappointed?

The auto-contouring marketing promise:

England will win the next Football World Cup

The product will select the most similar atlas
You will save more time
Atlas based auto-contouring

Single-atlas auto-contouring

Multi-atlas auto-contouring

Selection of single atlas

Selection of multiple atlases

DIR and contour warping

Contour consensus

Patient image

Patient image

Patient image

Patient image
Can atlas selection make results better?

Hypothesis 1: Using an atlas more similar to the patient will have a better contouring performance compared to the use of a randomly chosen atlas.

But how much better?

316 Head & Neck
Contours + Image

Patient image
Contours

Evaluate performance
- Dice Similarity Coefficient (DSC)
- Hausdorff Distance (HD)
- Average Distance (AD)
- Root Mean Square Distance (RMSD)

99540 atlas-patient pairs
Extreme Value Theory

An example from Oxford – Once-in-a-hundred-year flooding:

Generalised Pareto Distribution
Extreme Value Theory

- Fit GPD model to extreme cases

- Estimate the magnitude *expected* of one-in-a-hundred-years rainfall: 194 mm

- Estimate the achievable auto-contouring performance for a database of 5000 atlases
Multi-atlas results

- Estimated *expected* performance assuming an atlas database of 5000 atlases
- Contour fusion results assumes 10 best

<table>
<thead>
<tr>
<th>Structure</th>
<th>Single Atlas - DSC</th>
<th>Contour Fusion - DSC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Best</td>
<td>Est.</td>
</tr>
<tr>
<td>Brain</td>
<td>0.988</td>
<td>0.996</td>
</tr>
<tr>
<td>Brainstem</td>
<td>0.903</td>
<td>0.943</td>
</tr>
<tr>
<td>Oral Cavity</td>
<td>0.911</td>
<td>0.909</td>
</tr>
<tr>
<td>Parotid L</td>
<td>0.880</td>
<td>0.920</td>
</tr>
<tr>
<td>Parotid R</td>
<td>0.877</td>
<td>0.917</td>
</tr>
<tr>
<td>Spinal cord</td>
<td>0.891</td>
<td>0.887</td>
</tr>
<tr>
<td>Submandibular gl. L</td>
<td>0.873</td>
<td>0.872</td>
</tr>
<tr>
<td>Submandibular gl. R</td>
<td>0.863</td>
<td>0.862</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Metric</th>
<th>Single Atlas</th>
<th>Contour Fusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>HD (mm)</td>
<td>4.54</td>
<td>4.61</td>
</tr>
<tr>
<td>AD (mm)</td>
<td>0.97</td>
<td>0.98</td>
</tr>
<tr>
<td>RMSD (mm)</td>
<td>1.28</td>
<td>1.29</td>
</tr>
</tbody>
</table>
Good news!

With the most similar atlas
You will save more time
The auto-contouring will be better
The problem

Estimates the potential performance for a very large atlas database...

...assuming perfect atlas selection by an oracle with fore-knowledge of the output performance
Can we select the best atlas?

**Hypothesis 2:** Using an atlas with an image more similar to the patient image will have a better contouring performance compared to the use of a randomly chosen atlas.

Let’s test this

Extensive comparison of methods in the literature and commercially available

- Local and global descriptors
- Direct atlas search and template-based search
- 154 selection methods were implemented

Intensity (SSD, NMI, ...)
Deformation (Jacobian, ...)
Tissue overlap (bones, ...)

\[ N_A \text{ registrations} \]

\[ N_T \text{ registrations} \]
Performance analysis: Atlas selection

Results

Brainstem: DICE

Right Parotid: DICE

Online atlas selection methods
34 out of 154 shown
Performance analysis: Atlas selection

Results

Box plots of results on the Right Parotid
Why doesn’t atlas selection work?
Why doesn’t atlas selection work?

- Poor correlation of similarity measures to meaningful image differences
- Poor correlation image differences to meaningful contour differences
- Image measures can’t account for Contouring variability
The end result

Over enthusiastic marketing
Untested assumptions

Expectation ≠ Reality
<table>
<thead>
<tr>
<th>Reference</th>
<th>Anatomy</th>
<th>No. of Patients</th>
<th>No. of Atlases</th>
<th>Time saved (mins)</th>
<th>Time saved (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>segmentation of multiple target volumes and normal tissue (swallowing/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mastication) structures in the head and neck.&quot; <em>International Journal</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>of Radiation Oncology</em> <em>Biology</em> <em>Physics</em> 81.4 (2011): 950-957.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Young, Amy V., et al. &quot;Atlas-based segmentation improves consistency</td>
<td>Edometrial</td>
<td>10</td>
<td>15</td>
<td>9</td>
<td>26</td>
</tr>
<tr>
<td>and decreases time required for contouring postoperative endometrial</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cancer nodal volumes.&quot; <em>International Journal of Radiation Oncology</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Biology</em> <em>Physics</em> 79.3 (2011): 943-947.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gambacorta, Maria Antonietta, et al. &quot;Clinical validation of atlas-</td>
<td>Rectal</td>
<td>10</td>
<td>4</td>
<td>13</td>
<td>34</td>
</tr>
<tr>
<td>based auto-segmentation of pelvic volumes and normal tissue in rectal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tumors using auto-segmentation computed system.&quot; <em>Acta Oncologica</em> 52.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hwee, Jeremiah, et al. &quot;Technology assessment of automated atlas based</td>
<td>Prostate</td>
<td>5</td>
<td>75</td>
<td>3</td>
<td>24</td>
</tr>
<tr>
<td>s in the clinical workflow.&quot; *International Journal of Radiation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oncology* <em>Biology</em> <em>Physics</em> 72.1 (2008): S328-S329.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Granberg, Christoffer. &quot;Clinical evaluation of atlas based segmentation</td>
<td>Prostate</td>
<td>10</td>
<td>15</td>
<td>10</td>
<td>26</td>
</tr>
<tr>
<td>Langmack, K. A., et al. &quot;The utility of atlas-assisted segmentation in</td>
<td>Prostate</td>
<td>11</td>
<td>8</td>
<td>10</td>
<td>40</td>
</tr>
<tr>
<td>the male pelvis is dependent on the interobserver agreement of the</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>atlas-based organ-at-risk auto-segmentation-assisted radiation planning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>in head and neck cancer.&quot; <em>Radiotherapy and Oncology</em> 112.3 (2014):</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>321-325.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Take home messages – Part 1

⚠️ Question the assumption made in marketing

⚠️ Consider the workflow impact
Multi-modal target contouring
Multi-modal target contouring
Multi-modal target contouring
Where are we on the adoption curve?

Does your institution **routinely use** multi-modal target contouring?

A. Yes – PET/CT only
B. Yes – MRI only
C. Yes – PET/CT and MRI
D. No
E. Don’t know

http://aamd2017.participoll.com
Where are we on the adoption curve?

Multi-modal target contouring

- Wide range of solutions
- Clinical use standardising
- Vendors competing on features / quality

2.5% INNOVATORS
13.5% EARLY ADOPTERS
34% EARLY MAJORITY ADOPTERS
34% LATE MAJORITY ADOPTERS
16% LAGGARDS
PET/CT

✅ Shows metabolically active areas

⚠️ Not anatomical

⚠️ Low resolution
MRI

- Better soft tissue contrast
- Large range in acquisitions
- Often low res out of plane
- Limited field of view
- May be off-axis
- May have geometric distortion
How well integrated is it?

Where does your institution do multi-modal target contouring/registration?

A. Yes – all integrated on the TPS
B. Yes - 3rd Party
C. N/A – We don’t do it
D. Don’t know

http://aamd2017.participoll.com
It’s about improving treatment

Multi-modal contouring will improve treatment
Work without Multi-modal target contouring

Simulation
CT

Dosimetrist
OAR contouring
Planning

Rad Onc
Target contouring
Plan approval

Med Phys
Plan Check

Treatment

©Copyright Mirada Medical 2017
So why can it feel painful?

Dosimetrist

Simulation CT

Get MRI or PET

Registration

OAR contouring

Planning

Rad Onc

Target contouring

Plan approval

Med Phys

Registration

Registration QA

Plan check

Treatment

©Copyright Mirada Medical 2017
Let’s talk about efficiency

The ratio of the useful work performed by a machine or in a process to the total energy expended or heat taken in.

Photo: Raphael Matos, InsideTheRace.com
How is multi-modal target contouring increasing efficiency?

Simulation
CT

Get MRI or PET

Registration

OAR contouring

Planning

Dosimetrist

Rad Onc

Med Phys

Useful work + extra effort

Plan check

Target contouring

Plan approval

Registration

Registration QA

Treatment

Monitoring

CT

CT

CT
How do you do it?

How does your institution perform multi-modal target contouring?

A: Yes – Rigid registration only
B: Yes – Deformable registration
C: No – We still don’t use it!
D: Don’t know

http://aamd2017.participoll.com
Image registration

• Similarity measure
  – Performs anatomical matching

• Transformation model
  – Constrains the solution

• Regularization
  – Prevents matching noise
# Common similarity measures

<table>
<thead>
<tr>
<th>Similarity measure</th>
<th>Assumption</th>
<th>PET/CT</th>
<th>MRI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sum-Squared-difference (SSD)</td>
<td>Intensities in both images are the same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td>Intensities not the same, but have same order. i.e. brighter in one image = brighter in the other image</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modelled-SSD</td>
<td>Simple linear modelling can be used to map intensities. SSD performed on mapped intensities</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mutual Information (MI)</td>
<td>Some relationship between intensities in each image</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
SSD

Intensity in 1st image vs. Intensity in 2nd image

SSD
### Common similarity measures

<table>
<thead>
<tr>
<th>Similarity measure</th>
<th>Assumption</th>
<th>PET/CT</th>
<th>MRI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sum-Squared-difference (SSD)</td>
<td>Intensities in both images are the same</td>
<td>✔</td>
<td>!</td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modelled-SSD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mutual Information (MI)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Correlation

Intensity in 1st image vs. Intensity in 2nd image

- SSD
- CC

Graph showing correlation between intensity in the first and second images.
## Common similarity measures

<table>
<thead>
<tr>
<th>Similarity measure</th>
<th>Assumption</th>
<th>PET/CT</th>
<th>MRI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sum-Squared-difference (SSD)</td>
<td>Intensities in both images are the same</td>
<td>✔️</td>
<td>!</td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td>Intensities not the same, but have same ordering. i.e. brighter in one image = brighter in the other image</td>
<td>✔️</td>
<td>!</td>
</tr>
<tr>
<td>Modelled-SSD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mutual Information (MI)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Modelled SSD

Intensity in 1st image vs. Intensity in 2nd image

©Copyright Mirada Medical 2017
## Common similarity measures

<table>
<thead>
<tr>
<th>Similarity measure</th>
<th>Assumption</th>
<th>PET/CT</th>
<th>MRI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sum-Squared-difference (SSD)</td>
<td>Intensities in both images are the same</td>
<td>✓</td>
<td>!</td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td>Intensities not the same, but have same ordering. i.e. brighter in one image = brighter in the other image</td>
<td>✓</td>
<td>!</td>
</tr>
<tr>
<td>Modelled-SSD</td>
<td>Simple prior modelling can be used to map intensities. SSD performed on mapped intensities</td>
<td>✓ !</td>
<td>! !</td>
</tr>
<tr>
<td>Mutual Information (MI)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Common similarity measures

<table>
<thead>
<tr>
<th>Similarity measure</th>
<th>Assumption</th>
<th>PET/CT</th>
<th>MRI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sum-Squared-difference (SSD)</td>
<td>Intensities in images match</td>
<td>✔️</td>
<td>❗️</td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td>Intensities not the same, but have same order. i.e. brighter in one image = brighter in the other image</td>
<td>✔️</td>
<td>❗️</td>
</tr>
<tr>
<td>Modelled-SSD</td>
<td>Simple linear modelling can be used to map intensities. SSD performed on mapped intensities</td>
<td>✔️❗️</td>
<td>❗️</td>
</tr>
<tr>
<td>Mutual Information (MI)</td>
<td>Some relationship between intensities in each image</td>
<td>✔️❗️</td>
<td>✔️</td>
</tr>
</tbody>
</table>
Transformation models

Constrains the solution to a smaller range of possibilities
Rigid registration
Free-form deformable
## B-Splines

![B-Splines Diagram](image-url)
Radial Basis Functions
Transformation models

Constrains the solution to a smaller range of possibilities

✅ More constrained ➔ Faster (potentially)

⚠️ More constrained ➔ Smoother

⚠️ More constrained ➔ Less accurate
Regularizers

Prevent the similarity measure over fitting to noise
Regularizers

Similarity of images

Smoothness of deformation
Regularizers

Prevent the similarity measure over fitting to noise

- More constrained → Faster (potentially)
- More constrained → Smoother
- More constrained → Less accurate
So what’s all this go to do with efficiency?

The ratio of the useful work performed by a machine or in a process to the total energy expended or heat taken in.

Photo: Raphael Matos, InsideTheRace.com
Reducing efficiency

Wrong similarity measure

Over constrained

Under constrained

Less useful work

Lower efficiency
Using the registration

⚠️ Resampling loses information
Using the registration
Efficiency choice

- Contouring
- Resample contour
- Target contoured
- Resample image
- Contouring
- Registration

Harder to integrate
Maintains accuracy
Loss of accuracy
Easier to integrate
Take home messages – part 2

⚠️ No gain without pain!

⚠️ Be wary of resampling the image
Adaptive Therapy

Acquiring a new planning CT mid-treatment to account for changes in patient anatomy / response to treatment
Where are we on the adoption curve?

What percentage of your patients do you routinely consider adaptive therapy for?

A  None
B  1 – 10 %
C  10 – 30 %
D  30 – 50 %
E  50 – 100 %

http://aamd2017.participoll.com
Where are we on the adoption curve?

What percentage of your patients do you routinely **do** adaptive therapy for?

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>None</td>
<td>B</td>
<td>1 – 5 %</td>
<td>C</td>
</tr>
<tr>
<td>D</td>
<td>10 – 20 %</td>
<td>E</td>
<td>20 – 40 %</td>
<td>F</td>
</tr>
</tbody>
</table>

http://aamd2017.participoll.com
Where are we on the adoption curve?

Adaptive therapy

- 2.5% INNOVATORS
- 13.5% EARLY ADOPTERS
- 34% EARLY MAJORITY ADOPTERS
- 34% LATE MAJORITY ADOPTERS
- 16% LAGGARDS
Where are we really on the adoption curve?

What does your institution do before deciding to adapt?

A. Physical assessment of patient by the Rad Onc
B. Visual assessment of CBCT
C. Quantitative structure assessment on CBCT
D. Quantitative dose assessment on CBCT without recalculation
E. Quantitative dose assessment on CBCT with dose recalculation
F. Quantitative dose assessment on re-planning CT

http://aamd2017.participoll.com
Where are we really on the adoption curve?

What does your institution do for adaptive therapy?

A. Replan on new CT, without dose volume summation
B. Replan on new CT, using rigid dose from previous plan
C. Replan on new CT, using deformed dose from previous plan
D. Replan on new CT, using cumulative dose from previous plan across all CBCTs

http://aamd2017.participoll.com
Where are we on the adoption curve?

Adaptive therapy

- Multiple components needed
- Clinical benefit being evaluated
- Vendors need to collaborate with Innovators and each other

Increasing effort

2.5% INNOVATORS 13.5% EARLY ADOPTERS 34% EARLY MAJORITY ADOPTERS 34% LATE MAJORITY ADOPTERS 16% LAGGARDS
Adaptive Therapy – An efficiency choice

What would the increase in quality can I achieve if ...?  How much time can I spend/tolerate?

Treatment quality  Effort/time required
Adaptive therapy is a broad spectrum

We support adaptive therapy

You can improve quality
Take home message 3

⚠️ No gain without pain!

⚠️ Clinical evidence will drive efficiency

⚠️ Clinical evidence will drive software effort reduction
Take home messages – the summary

⚠️ Question the assumption made in marketing

⚠️ No gain without pain!

⚠️ Clinical evidence will drive software effort reduction