Deformable Image Registration in Radiation Therapy

Kristy K. Brock, PhD, DABR, FAAPM
Professor
Department of Imaging Physics
Department of Therapy Physics
University of Texas MD Anderson Cancer Center

Acknowledgements & Conflict of Interest

• AAPM Task Group 132 Chair
• I have a licensing agreement for deformable image registration technology with RaySearch Laboratories.
• I recently held a co-development agreement with Varian Medical Systems
• Currently hold a research agreement with RaySearch Laboratories
• AAPM TG 132
• Molly McCulloch, PhD, Guillaume Cazoulat PhD, Peter Balter, PhD, Andrea Ohrt
• Clifton D. Fuller, MD, PhD – MDACC
• Morfeus Lab – past and present
• Fuller Lab

Motivation

• Numerous retrospective, limited data studies showing that planned ≠ delivered dose and there is an improvement in therapeutic ratio with adaptation
  – Prospective HN trial – all patients had 1 re-plan, 33% had 2 re-plans
• Growing demand from the radiotherapy community – from large academic centers to single vault community practitioners – to have efficient, easy to use tools available
• Increasing number of RTOG/NRG trials including adaptation
Motivation

- Clinical Goal: automatically accumulate dose on all patients, use this data to design intelligent adaptation guidelines, and design clinical trials to reduce toxicity/improve tumor control using adaptive (anatomical and functional) tools
- Integration and automation of the tools is essential
  - Even one break in the chain or missing tool can break the whole workflow.
SAFETY!
Commissioning and QA
Recommendations from TG 132

First: A Word of Caution

Example: Multi-modality imaging for Planning
Liver: CT (No Contrast = No visible GTV)
Liver: MR (Visible GTV)
Clinical Registration
X: 26.1mm Y: 119.8mm Z: -12.6mm
X: 1.9deg Y: -2.9deg Z: -4.6deg

Auto, liver last step
X: 25.6mm Y: 120.8mm Z: -26.1mm
X: -1.5deg Y: 2.5deg Z: -3.4deg

Nearby Structure Map
X: 14.5mm Y: 122.3mm Z: -26.1mm
X: -1.5deg Y: 2.5deg Z: 4.1deg
Liver Contour optimization
X: 13.0mm Y: 125.3mm Z: -19.0mm
X: 0.4deg Y: -1.3deg Z: 2.3deg

Overall Comparison [mm, Degrees]

<table>
<thead>
<tr>
<th>Registratio n</th>
<th>dX</th>
<th>dY</th>
<th>dZ</th>
<th>Xrots</th>
<th>Yrots</th>
<th>Zrots</th>
<th>Overlap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical</td>
<td>26.1</td>
<td>119.8</td>
<td>12.6</td>
<td>1.9</td>
<td>-2.9</td>
<td>-4.6</td>
<td>Defined</td>
</tr>
<tr>
<td>Auto</td>
<td>25.6</td>
<td>120.8</td>
<td>-26.1</td>
<td>-1.5</td>
<td>2.5</td>
<td>-3.4</td>
<td></td>
</tr>
<tr>
<td>Vessel</td>
<td>14.5</td>
<td>122.3</td>
<td>26.1</td>
<td>-1.5</td>
<td>2.5</td>
<td>4.1</td>
<td></td>
</tr>
<tr>
<td>Boundary</td>
<td>13.0</td>
<td>125.3</td>
<td>19.0</td>
<td>0.4</td>
<td>-1.3</td>
<td>2.3</td>
<td></td>
</tr>
</tbody>
</table>

Image Registration: Accurate Target Definition
Clinical Effect

GTV (defined on MR, mapped to CT for Tx)

- Assess uncertainty around GTV
- Add margin around GTV definition to account for uncertainty when required

Region of CT-defined GTV that is missed

Clinical Recommendations (1/2)

1. Understand the basic image registration techniques and methods of visualizing image fusion
2. Understand the basic components of the registration algorithm used clinically to ensure its proper use
3. Perform end-to-end tests of imaging, registration, and planning/treatment systems if image registration is performed on a stand-alone system

Clinical Recommendations (2/2)

4. Perform comprehensive commissioning of image registration using the provided digital phantom data (or similar data) as well as clinical data from the user’s institution
5. Develop a request and report system to ensure communication and documentation between all users of image registration
6. Establish a patient specific QA practice for efficient evaluation of image registration results
Commissioning and QA
Understand the whole picture

How do they work?

• Match something
  – Intensity, gradients, boundaries, features
• Constrain by a function
  – Geometric, physical, biomechanical

How do they work?

• Match something
  – Intensity, gradients, boundaries, features
  – What happens when the intensity correspondence varies?
  – What happens when the gradient isn’t there?
  – What happens when the boundaries aren’t well defined?
  – What happens with the features aren’t visible?
• Constrain by a function
  – Geometric, physical, biomechanical
  – Can you rely on this model when the match above is missing?
Great Car! Right?
*Ferrari California T*

Is it a great car for a road trip?

*It is NOT a great car for THIS application!*

Sum of Squared Differences

\[ \sum_{C1}^{C2} (I_{C1} - I_{C2})^2 \]

... subtract one image from the other

**Individual Intensity Distributions**

**Sum of the Squares of the Differences**
This doesn’t usually make much sense

\[ I_{CT} - I_{MR} \neq 0 \]

Individual Intensity Distributions

\[ I_{CT} \]

\[ I_{MR} \]

\[ \text{Differene Image} \]

\[ \text{Not Zero} \]

\[ \text{This doesn’t usually make much sense} \]

How Reliable is the Max MI?

- Actually, min -MI

\[ \text{Min -MI} \]

- Best Solution

\[ \text{Min -MI} \]

- Best Solution

Intensity Variation: Impact on CC/MSD

Clear intensity variation

No relevant intensity variation, noise/artifact
Understand the basic components of the registration algorithm used clinically to ensure its proper use

How?

- At minimum, the vendor should disclose:
  - Similarity metric used
  - Regularization used
  - Transformation used
  - Optimization method
  - What knobs you can turn and what they do
- Read white papers

Why do we need to know the implementation?

Med Phys 2008

New method to validate Deformable Image Registration

Deformable 3D Presage dosimeters

Control (No Deformation)  Deformed (27% Lateral Compression)

Slides Courtesy of Mark Oldham and Shiva Das
New method to validate Deformable Image Registration

CT of optical 3D dosimeter

Dose Map Dose

Slides Courtesy of Mark Oldham and Shiva Das

Dosimeter & Deformable Registration-based Dose Accumulation: Dose Distributions

Field Shape Differences

Caution must be used when accumulating dose, especially in regions of the image with homogeneous intensity.

Distribution

Pre-Processing Prior to DIR

Measured, Optical CT

Requires little/no pre-processing

DIR-predicted, Intensity-based DIR

Requires contours of structures

DIR-predicted, Biomechanical

Surface projection

Slides Courtesy of Mark Oldham and Shiva Das

Different DIR Algorithms have Different Strengths and Weaknesses

### Different DIR Algorithms have Different Strengths and Weaknesses

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Coronal</th>
<th>Axial</th>
<th>Sagittal</th>
<th>3D Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measured, Optical CT</td>
<td>![Image]</td>
<td>![Image]</td>
<td>![Image]</td>
<td>96% (^1) (control)</td>
</tr>
<tr>
<td>DIR-predicted, Intensity-based DIR</td>
<td>![Image]</td>
<td>![Image]</td>
<td>![Image]</td>
<td>60% (^1)</td>
</tr>
<tr>
<td>DIR-predicted, Biomechanical Surface projection</td>
<td>![Image]</td>
<td>![Image]</td>
<td>![Image]</td>
<td>91% (^2)</td>
</tr>
</tbody>
</table>


---

### Commissioning and QA

**Understand the whole picture**

- Phantom approach to understand characteristics of algorithm implementation
- Complex algorithm

---

### Conclusion

We evaluated the commercially available DDR software using coronal 3D-CT images from multiple centers. Our results demonstrated that DDR accuracy differed among institutions because it was dependent on both the DDR software and procedure. Our results could be helpful for establishing prospective clinical trials.
Rigid Geometric Data

- Helps us to learn the impact of the 'knobs' of the registration
- Validation of most straightforward case
- Similar to 20x20 field profile

* Phantom Data Courtesy of ImSim QA

Example Commissioning Tests

<table>
<thead>
<tr>
<th>Effect to Primary</th>
<th>±Δx</th>
<th>±Δy</th>
<th>±Δz</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.9543</td>
<td>-1.0</td>
<td>5.0</td>
<td>-1</td>
</tr>
<tr>
<td>0.9471</td>
<td>-1.0</td>
<td>5.0</td>
<td>-1</td>
</tr>
<tr>
<td>0.9543</td>
<td>-1.0</td>
<td>5.0</td>
<td>-1</td>
</tr>
<tr>
<td>0.9543</td>
<td>-1.0</td>
<td>5.0</td>
<td>-1</td>
</tr>
<tr>
<td>0.9543</td>
<td>-1.0</td>
<td>5.0</td>
<td>-1</td>
</tr>
<tr>
<td>0.9543</td>
<td>-1.0</td>
<td>5.0</td>
<td>-1</td>
</tr>
</tbody>
</table>

Rigid Anatomical Phantom

- Multi-Modality
- Translation Offset
- 1 additional (simple) layer of complexity
**Deformable Phantom**

**Commissioning Procedure:**
- Run Deformable Image Registration
- Export DICOM Deformation Vector Field (DVF)
- Pseudo code provided to compare known DVF with exported DVF
- Target: 95% of voxels within 2 mm, max error less than 5 mm

---

**Deformable Lung**

- Clinical Lung Data
- Simulated Deformed Lung

*Courtesy DIR-lab, Dr. Castillo*

---

**Commissioning and QA**

**Understand the whole picture**

- Phantom approach to understand characteristics of complex algorithms
- Quantitative validation of clinical images
Quantitative Validation Techniques

- **Landmark Based**
  - Does the registration map a landmark on Image A to the correct position on Image B?
  - Target Registration Error (TRE)
- **Contour Based**
  - Does the registration map the contours onto the new image correctly?
  - Dice Similarity Coefficient (DSC)
  - Mean Distance to Agreement (MDA)

Landmark Based (TRE)

- Reproducibility of point identification is sub-voxel
- Gross errors
- Quantification of local accuracy within the target
- Increasing the number increases the overall volume quantification
- Manual technique
- Can identify max errors

That sounds great! Is that enough?
Accuracy of Points

$$1 \text{ cm}$$

$$\bigcirc \quad \bigcirc \quad \bigcirc$$

RMS = 0.3 cm

Points Don’t Tell the Whole Story

Accuracy of Contours

Actual Exhale  Modeled Exhale  Iter  0.2 Bronchial Bits

Inhale  DSC > 0.9  Modeled Exhale  TRE: 8.0 mm

DSC > 0.9
**Request**

- Clear identification of the image set(s) to be registered
- Identification of the primary (e.g., reference) image geometry
- An understanding of the local region(s) of importance
- The intended use of the result
  - Target delineation
- Techniques to use (deformable or rigid)
- The accuracy required for the final use

---

### Quality Assurance

<table>
<thead>
<tr>
<th>Quality Assessed</th>
<th>Reason</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Whole scan aligned</td>
<td>- Hounsfield unit (HU) incorrect&lt;br&gt;- Inaccurate structure definition everywhere&lt;br&gt;- Image not of acceptable quality</td>
<td>- Inaccurate target definition&lt;br&gt;- Image not of acceptable quality</td>
</tr>
<tr>
<td>2 Locally aligned</td>
<td>- Hounsfield unit (HU) correct&lt;br&gt;- Image not of acceptable quality</td>
<td>- Inaccurate target definition&lt;br&gt;- Image not of acceptable quality</td>
</tr>
<tr>
<td>3 Usable with risk of deformation</td>
<td>- Hounsfield unit (HU) correct&lt;br&gt;- Image not of acceptable quality&lt;br&gt;- Additional information required</td>
<td>- Inaccurate target definition&lt;br&gt;- Image not of acceptable quality&lt;br&gt;- Additional information required</td>
</tr>
<tr>
<td>4 Usable for diagnosis only</td>
<td>- Registration not good enough to rely on parametric imaging&lt;br&gt;- Possible use to identify general location of lesion (e.g., PT for head)</td>
<td>- Inaccurate target definition&lt;br&gt;- Image not of acceptable quality&lt;br&gt;- Additional information required</td>
</tr>
<tr>
<td>5 Alignment not acceptable</td>
<td>- Unable to align anatomy to acceptable levels&lt;br&gt;- Hounsefield unit (HU) incorrect between scans (e.g., surgical contrast of the anatomy of interest or lesion not clearly visible)</td>
<td>- Inaccurate target definition&lt;br&gt;- Image not of acceptable quality&lt;br&gt;- Additional information required</td>
</tr>
</tbody>
</table>
Establish a patient specific QA practice for efficient evaluation of image registration results

Why?
• At this point we are still understanding how the registration is performing on different types of patients

How?
• Visual Verification
• Spot checks of landmarks
• Boundary comparison

Vendor Recommendations

1. Disclose basic components of their registration algorithm to ensure its proper use
2. Provide the ability to export the registration matrix or deformation vector field for validation
3. Provide tools to qualitatively evaluate the image registration
4. Provide the ability to identify landmarks on 2 images and calculate the TRE from the registration
5. Provide the ability to calculate the DSC and MDA between the contours defined on an image and the contours mapped to the image via image registration
6. Support the integration of a request and report system for image registration

TG-132 Product

• Guidelines for understating of clinical tools
• Digital (virtual) phantoms
• Recommendations for commissioning and clinical implementation
• Recommendations for periodic and patient specific QA/QC
• Recommendations for clinical processes
Clinical Examples

Transitioning DIR into the Clinic

- **Commissioning the DIR system in your clinic is important**
  - It will take resources
  - Similar to the start of IMRT, extra measurements will be needed as we begin to understand the tools
  - It will have a return on investment – Improved efficiency in the process
  - Important to commission the system in your clinic
- **Recognize the uncertainties**
- **Cost/benefit of Adaptive**
- **Design clinical workflow**

PI: CD Fuller MDACC
DIR Commissioning: CT to CT

- Repeat CTs
- Contours drawn on both
- DIR CT to CT
- Dice Similarity Coefficient (DSC) evaluation
  - Evaluates overlap of structures
  - Ranges from 0 (no overlap) to 1 (perfect overlap)
- Limited by manual contour reproducibility (~0.75-0.8)

<table>
<thead>
<tr>
<th>Structure</th>
<th>Dice Similarity Coefficient (DSC)</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brainstem</td>
<td>0.85</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>Larynx</td>
<td>0.81</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>R Parotid</td>
<td>0.78</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>L Parotid</td>
<td>0.77</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>R Submandibular Gland</td>
<td>0.77</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>L Submandibular Gland</td>
<td>0.78</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>Mandible</td>
<td>0.91</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>Esophagus</td>
<td>0.73</td>
<td>0.13</td>
<td></td>
</tr>
</tbody>
</table>

9 Patients, 2 CTs, Anaconda

DSC Differences in Practice

- DIR: planning CT to weekly adaptive CT
  - ANACONDA method generates a smooth DVF optimized by the quasi-Newton algorithm and guided by the correlation coefficient between the images.
- Map Contours based on registration
- Clinicians review contours and edit when needed
  - Editing is naturally driven by clinical importance
  - Dice Similarity Coefficient (DSC) has bias due to organ size and shape
DSC Difference between DIR-propagated Contours and Final Approved Clinician Contours

<table>
<thead>
<tr>
<th>Structure</th>
<th>DIR Angle</th>
<th>CT Angle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skull</td>
<td>1.00</td>
<td>0.99</td>
</tr>
<tr>
<td>Brainstem</td>
<td>0.96</td>
<td>0.99</td>
</tr>
<tr>
<td>Spinal Cord</td>
<td>0.98</td>
<td>0.97</td>
</tr>
<tr>
<td>Spinal Cord</td>
<td>0.97</td>
<td>0.97</td>
</tr>
<tr>
<td>Esophagus</td>
<td>0.96</td>
<td>0.96</td>
</tr>
<tr>
<td>Lung Submass</td>
<td>0.96</td>
<td>0.96</td>
</tr>
<tr>
<td>Oesophagus</td>
<td>0.96</td>
<td>0.96</td>
</tr>
<tr>
<td>Molars</td>
<td>0.94</td>
<td>0.95</td>
</tr>
<tr>
<td>Mandible</td>
<td>0.96</td>
<td>0.96</td>
</tr>
<tr>
<td>Medulla Oblongata</td>
<td>0.90</td>
<td>0.95</td>
</tr>
<tr>
<td>Soft Tissue</td>
<td>0.90</td>
<td>0.90</td>
</tr>
<tr>
<td>Brainstem</td>
<td>0.96</td>
<td>0.96</td>
</tr>
<tr>
<td>Spinal Cord</td>
<td>0.97</td>
<td>0.97</td>
</tr>
<tr>
<td>Spinal Cord</td>
<td>0.97</td>
<td>0.97</td>
</tr>
<tr>
<td>Esophagus</td>
<td>0.96</td>
<td>0.96</td>
</tr>
<tr>
<td>Lung Submass</td>
<td>0.96</td>
<td>0.96</td>
</tr>
<tr>
<td>Oesophagus</td>
<td>0.96</td>
<td>0.96</td>
</tr>
<tr>
<td>Molars</td>
<td>0.94</td>
<td>0.95</td>
</tr>
<tr>
<td>Mandible</td>
<td>0.96</td>
<td>0.96</td>
</tr>
<tr>
<td>Medulla Oblongata</td>
<td>0.90</td>
<td>0.95</td>
</tr>
<tr>
<td>Soft Tissue</td>
<td>0.90</td>
<td>0.90</td>
</tr>
</tbody>
</table>

What is the dosimetric impact?

DIR & Contour Propagation

External Beam Dose Calculation

CBCT Plan Evaluation: Dose Compare Each Fraction, Fx N

Planning CT Planned Dose Planning Fx N+1

Yes No

DIR Commissioning: CT to CBCT

- Getting contours on multiple CBCTs for validation is difficult & has larger uncertainties than CT
- Alternative: indirect validation
- Logic argument:
  - CT to CT DIR has been commissioned and deemed within contour uncertainty
  - CT and CBCTs obtained on the same day with immobilization should have minimal differences after rigid registration
Clinical Workflow Questions

- What is the difference between dose calculated on CT vs CBCT?
- Do we need daily CBCT dose accumulation or does weekly CT (or CBCT) give us the same answer?
- If we perform an adaptive replan, how do we accumulate the dose
  - Deform daily images back to planning image or all daily images back to the initial planning image?
  - Dose A to B to C equal A to C?
- Does the predictive model hold with true dose accumulation?
- How can we extrapolate the CBCT?
- How do we distill down the data
  - I have a 125 page document of data for 1 patient!!!
- However, we don’t need to have all of these answers to start!
Summary

• DIR is a powerful tool that can help us to integrate multi-modality images, understand motion and anatomical changes, and compute an improved estimate of the delivered dose.

• With power comes responsibility... we must commission the system prior to use, understand the limitations, and communicate its proper use to clinicians, dosimetrist, therapists, and others.

• The presentation of information leads to decision making... we must move use the data to design intelligent adaptation strategies to improve the therapeutic ratio efficiently.