DIR QA Options and Research Development

Neil Kirby, Ph.D.
Assistant Professor
Radiation Oncology
UTHSCSA
Acknowledgments

ICH: University of California, San Francisco
Jean Pouliot
Josephine Chen
Hojin Kim
Kamal Singhrao
Cynthia Chuang
Olivier Morin

ICH: University of Texas Health Science Center at San Antonio
Daniel Saenz
Kristen Cline
Mohammad Obeidat
Sotirios Stathakis

ICH: University of Florida Health Cancer Center – Orlando Health
Jason Pukala

ICH: Yale
Sarah Geneser

ICH: Rutgers
Ke Nie
Overview

• Phantom DIR QA.

• Lessons learned from phantoms.

• Patient-specific QA.
Types of DIR QA

Phantom
- Know the deformation.
- Not a real case.

Patient
- Limited knowledge of deformation.
- DIR errors matter.
Types of Phantoms

Physical

Before a)  After b)  2-D Deformation Field c)

Virtual

Physical Phantoms

Pros
• Are a full end-to-end QA of imaging and DIR.

Cons
• Requires specialized fabrication.
• May not fully represent a patient
How realistic should physical phantoms be?

Virtual Phantoms

Pros
• Closely resemble patients.
• Can be created digitally.
• Commercial platforms (ImSimQA from OSL).

Cons
• Not an end-to-end test.
Deformable Image Registration Evaluation Project (DIREP)

- Virtual phantoms are also available online: [https://sites.google.com/site/dirphantoms/](https://sites.google.com/site/dirphantoms/)

How realistic should virtual phantoms be?

- Image noise can leave a “fingerprint” of the deformation.
How realistic should virtual phantoms be?

<table>
<thead>
<tr>
<th>Scenario</th>
<th>MIM Mean Error [mm]</th>
<th>Velocity Mean Error [mm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Realistic Noise Scenario</td>
<td>1.24</td>
<td>1.68</td>
</tr>
<tr>
<td>No Processing</td>
<td>0.78</td>
<td>1.66</td>
</tr>
</tbody>
</table>

• This fingerprint affects some algorithms more than others.
Virtual phantoms to represent daily imaging

How realistic should virtual phantoms be?

Virtual phantoms must include:

• 1. Realistic noise scenarios.

• 2. Realistic deformations – missing tissue, divergences, and discontinuities in the deformation field.
Virtual phantoms can be

<table>
<thead>
<tr>
<th>Percentage</th>
<th>Option</th>
</tr>
</thead>
<tbody>
<tr>
<td>20%</td>
<td>1. downloaded from online.</td>
</tr>
<tr>
<td>20%</td>
<td>2. created with vendor software.</td>
</tr>
<tr>
<td>20%</td>
<td>3. useful for DIR QA.</td>
</tr>
<tr>
<td>20%</td>
<td>4. all of the above.</td>
</tr>
<tr>
<td>20%</td>
<td>5. none of the above.</td>
</tr>
</tbody>
</table>

10
Some Basic Types of DIR QA

1. Contour mapping.
2. Landmark Tracking.
3. Image similarity.
4. Comparison of predictions to known deformation (for phantoms).
Contour Mapping

•Comparison of transferred contours to that drawn on the other image.
• Landmarks are identified on the CT images and tracked to test deformation predictions.

• Any prominent landmark will also stand out to the algorithm.
A mathematical metric is used to evaluate the similarity between the warped image and the fixed image.

A few examples - sum of squared difference, mean absolute difference (MAD), cross correlation, and mutual information.

\[
MAD = \frac{1}{N} \sum_{i=1}^{N} |W_i - F_i|
\]
Balance of Penalty Terms

• Penalty Function = Similarity + λ • Regularization
• An optimal balance between the penalty terms must be found.

![Images showing the balance of penalty terms for different values of λ: Large, Medium, Small]
UCSF Deformable Pelvic Phantom

- Pelvic phantom to represent deformations from bladder filling.
- Realistic soft-tissue variations.
• MIM tested with a variable regularization factor (20 values).
• Overall spatial error (relevant for dose mapping) optimized with a different factor than contour transfer.
Comparison of QA Methods on Virtual Phantoms

• Similar results are found with virtual phantoms.

• **M. Obeidat** et al. “Comparison of Different QA Methods for Deformable Image Registration to the Known Errors for Prostate and Head-And-Neck Virtual Phantoms”, SU-E-J-117, Exhibit Hall
Lessons learned from phantoms

1. Contour mapping.
2. Landmark Tracking. Actual Accuracy.
3. Image similarity.
Reason for discrepancy

- A strong focus on image similarity tends to increase accuracy in regions of contrast, but decrease accuracy in homogenous regions.

- The 3 metrics are dominated by regions of contrast.
Lessons from phantoms

1. Contour mapping.
2. Landmark Tracking.
3. Image similarity.

Actual Accuracy.

Available for Patients
DIR QA for a patient.

• For some applications, these metrics are adequate for DIR QA.

• For contour transfer, you simply need to visually inspect these contours.
DIR QA for a patient.

• For other applications, such as dose or standard uptake values, these metrics are not enough.

• These are deposited in both high and low contrast regions.
DIR QA for a patient.

- Evaluating DIR spatial accuracy must be broken into two parts:

  1. Accuracy assessment in regions of contrast.

  2. Evaluation of how physical the deformation field is away from regions of contrast

  (Jacobian, vector field, deformation grid)
• You can estimate the spatial accuracy from these two terms and multiply by local dose gradient.

2.5 Gy/mm x 1.3 mm = 3.25 Gy

• You should still assume at least a 1 mm error.

• B. Tewfik et al. “Submillimeter accuracy in radiosurgery is not possible”, Med Phys, 40, 050601 (2013)
Automated DIR QA for a patient.

- Manual techniques for quantitative DIR accuracy assessment are not incredibly practical.

- Automated software to create DIR mapping uncertainty is the next step for DIR QA.
Automated DIR QA for a patient.

Overall spatial accuracy of a DIR algorithm can be directly evaluated by its ability to:

1. transfer anatomical contours. 20%
2. deform landmarks. 20%
3. produce close image similarity. 20%
4. all of the above. 20%
5. none of the above. 20%
Conclusions

• How realistic a phantom (virtual or physical) is will affect DIR accuracy.

• DIR effectiveness for contour mapping, landmark tracking, and image similarity is not the same as overall spatial accuracy.

• Patient QA should assess both image similarity and how physical a deformation is.