DIR QA Options and Research Development

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

UT HEALTH SCIENCE CENTER®
SAN ANTONIO

Acknowledgments

UTHSCSA

Daniel Saenz

Kristen Cline

Mohammad Obeidat

Sotirios Stathakis

UF Health Cancer Center – Orlando Health

Jason Pukala

Yale

Sarah Geneser

UCSF

Jean Pouliot

Josephine Chen

Hojin Kim

Kamal Singhrao

Cynthia Chuang

Olivier Morin

Rutgers

Ke Nie

Overview

- Phantom DIR QA.
- Lessons learned from phantoms.
- Patient-specific QA.

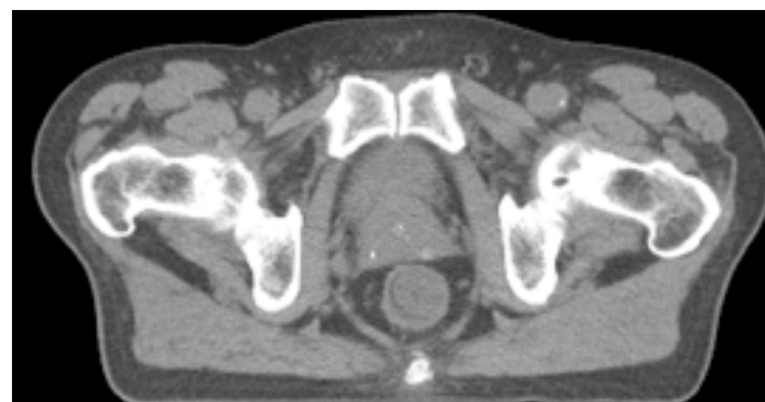
Types of DIR QA

Phantom

- Know the deformation.
- Not a real case.

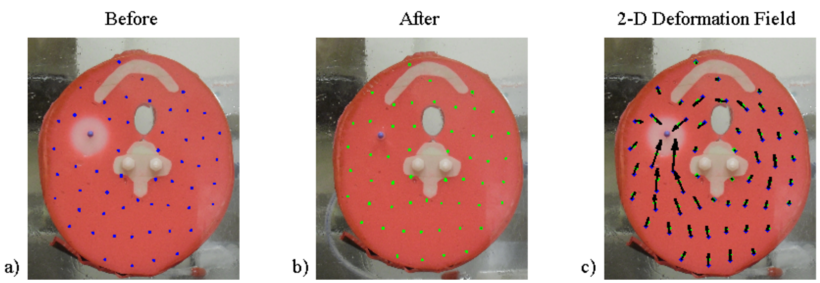
Patient

- Limited knowledge of deformation.
- DIR errors matter.

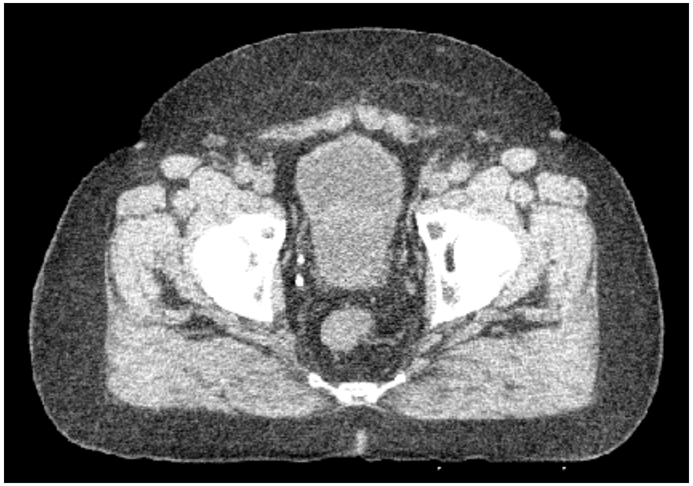


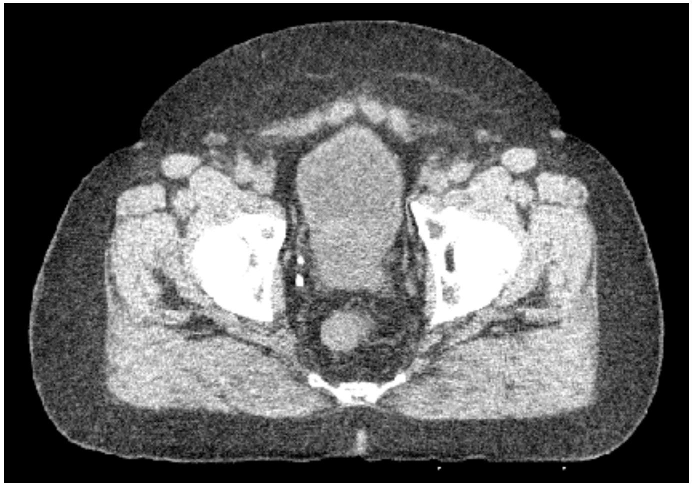
Types of Phantoms

Physical



Virtual





K. Nie et al. Med Phys **40**, 041911 (2013).

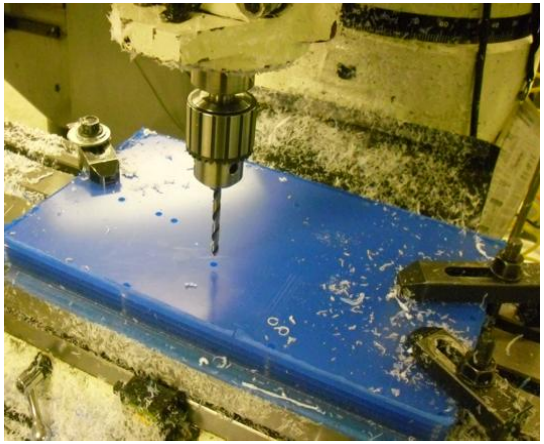
Physical Phantoms

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?

- **D. Saenz** et al. “Quality Assurance of Deformable Image Registration Algorithms: How Realistic Should Phantoms Be?”,

  SU-E-J-97, Exhibit Hall

-

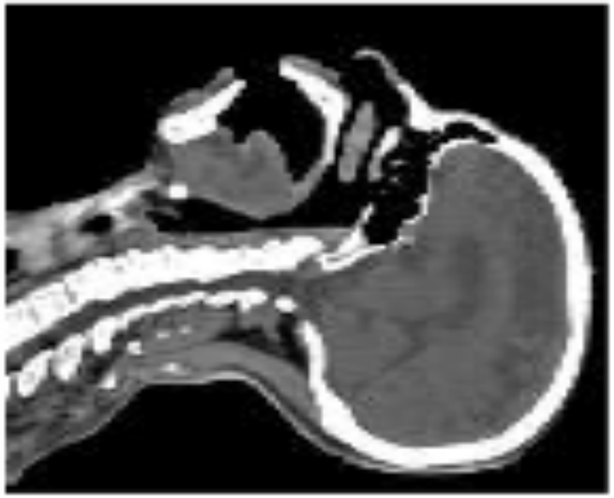
Virtual Phantoms

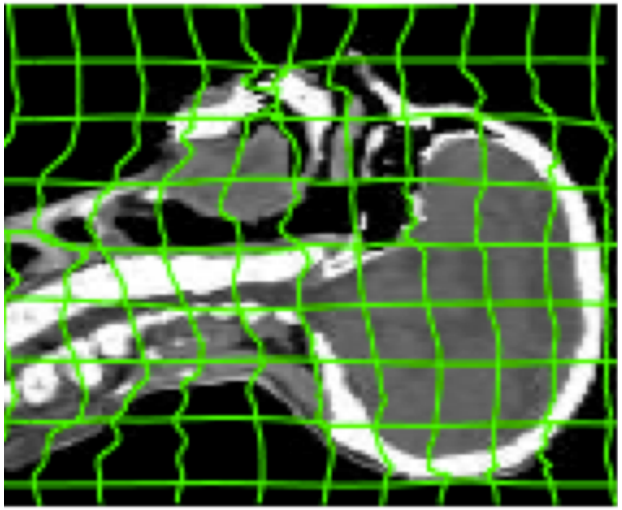
Pros

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

Cons

- Not an end-to-end test.



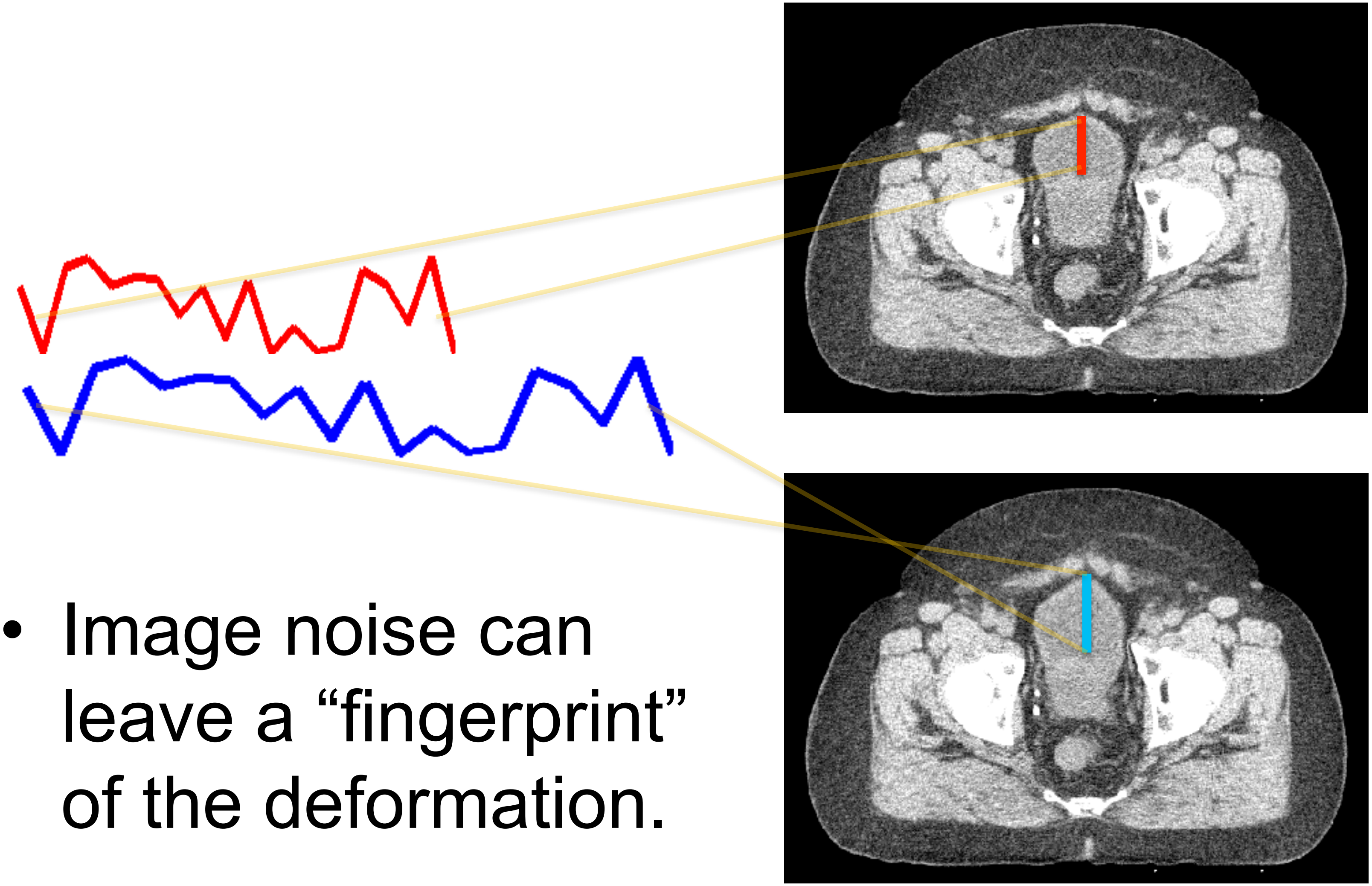


Deformable Image Registration Evaluation Project (DIREP)

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

J. Pukala et al. Med Phys **40**, 111703 (2013).

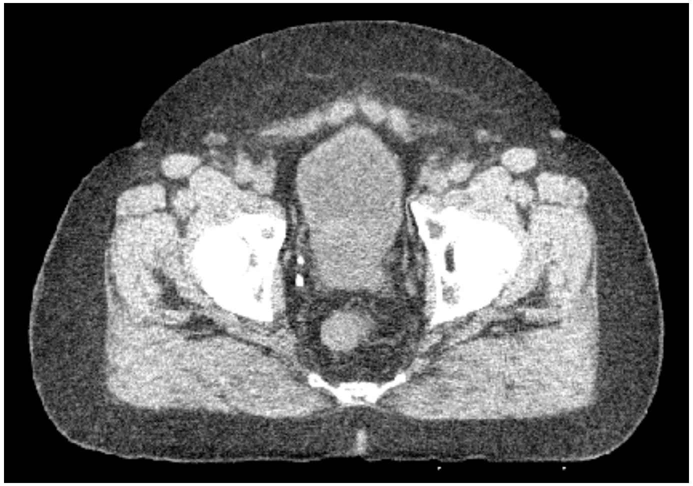
How realistic should virtual phantoms be?



- Image noise can leave a "fingerprint" of the deformation.

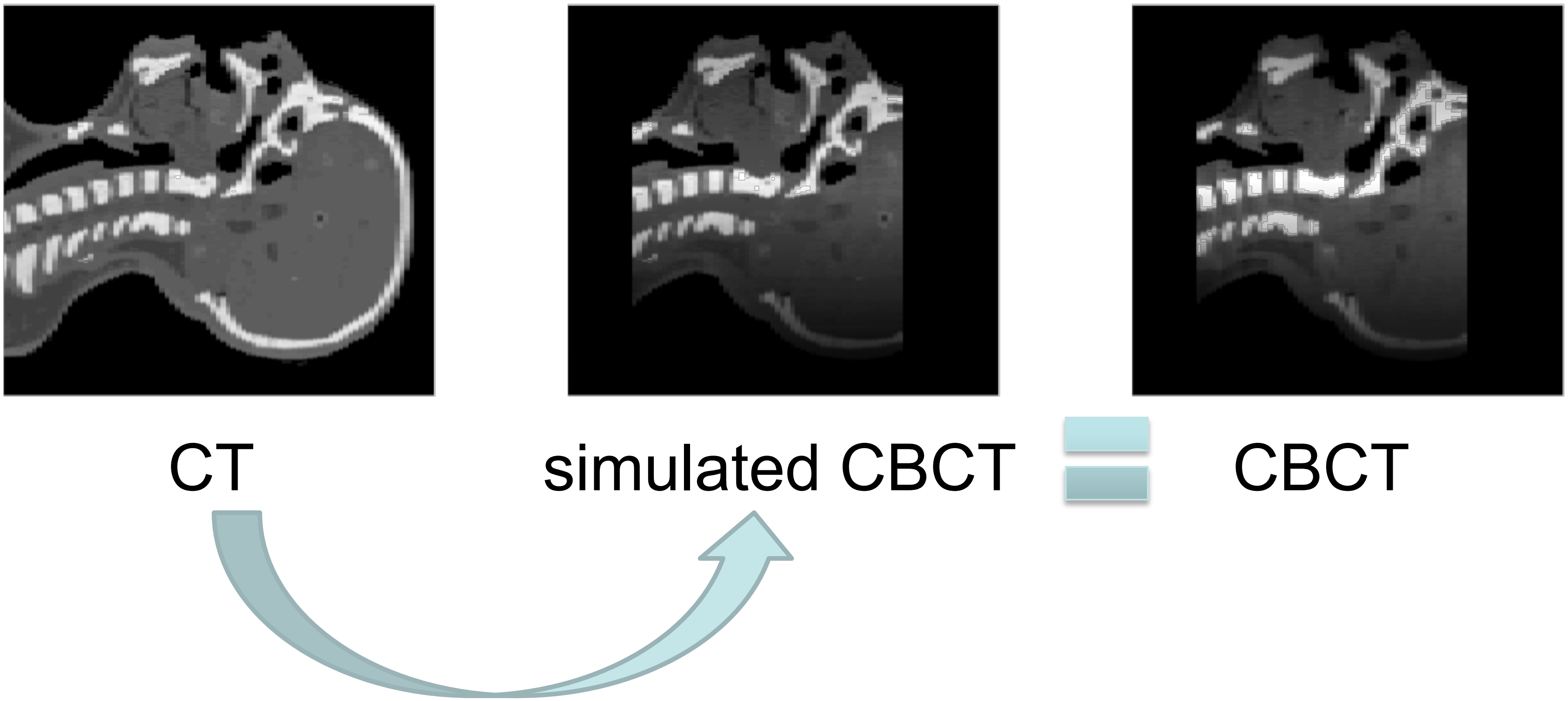
How realistic should virtual phantoms be?

	MIM Mean Error [mm]	Velocity Mean Error [mm]
Realistic Noise Scenario	1.24	1.68
No Processing	0.78	1.66



- This fingerprint affects some algorithms more than others.

Virtual phantoms to represent daily imaging



- **K. Cline** et al. “Comparative Analysis of MIM and Velocity's Image Deformation Algorithm Using Simulated KV-CBCT Images for Quality Assurance

SU-E-J-89, Exhibit Hall

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.

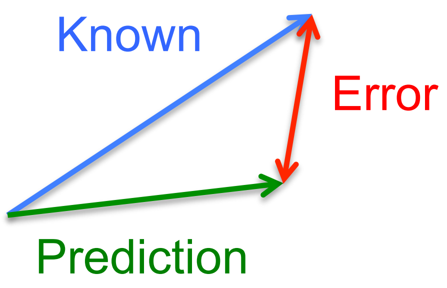
Virtual phantoms can be

20%	1.	downloaded from online.
20%	2.	created with vendor software.
20%	3.	useful for DIR QA.
20%	4.	all of the above.
20%	5.	none of the above.

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.

Landmark Tracking

- Landmarks are identified on the CT images and tracked to test deformation predictions.

- Any prominent landmark will also stand out to the algorithm.

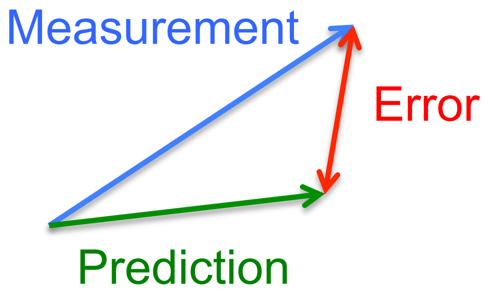
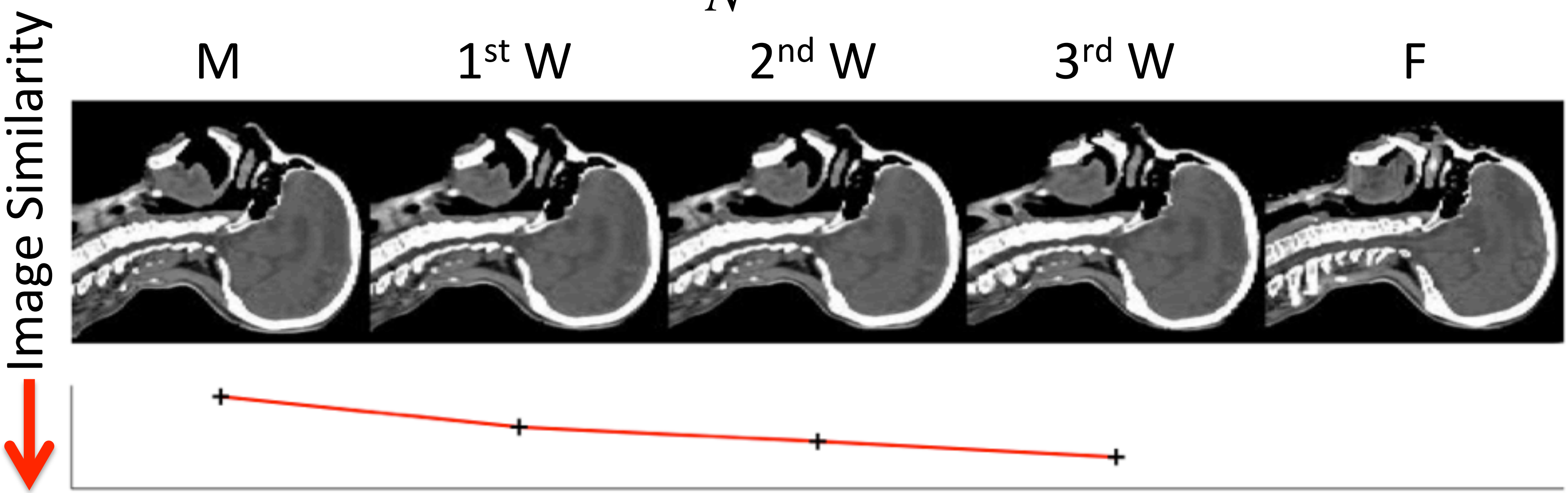


Image Similarity

- A mathematical metric is used to evaluate the similarity between the warped image and the fixed image.

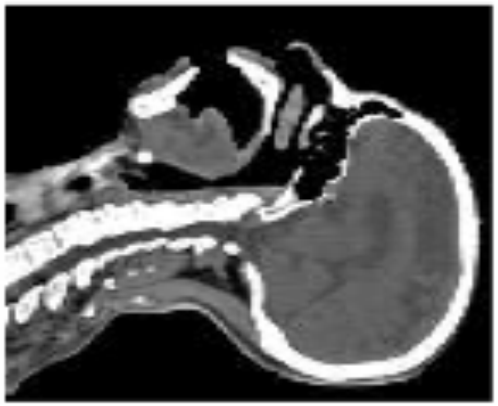
- A few examples - sum of squared difference, mean absolute difference (MAD), cross correlation, and mutual information.

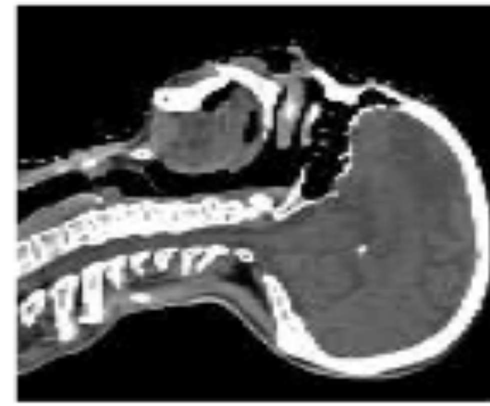
$$MAD = \frac{1}{N}\sum_{i=1}^{N}\left|W_i - F_i\right|$$



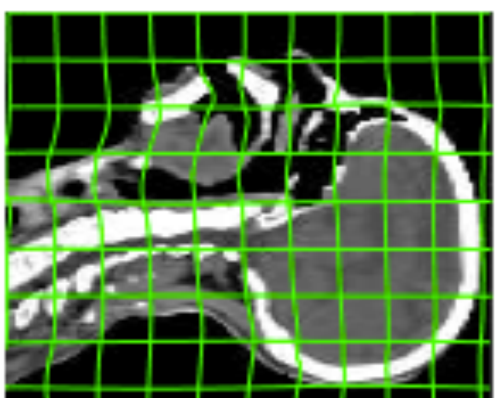
Balance of Penalty Terms

- Penalty Function = Similarity + $\lambda \cdot$ Regularization
- An optimal balance between the penalty terms must be found.

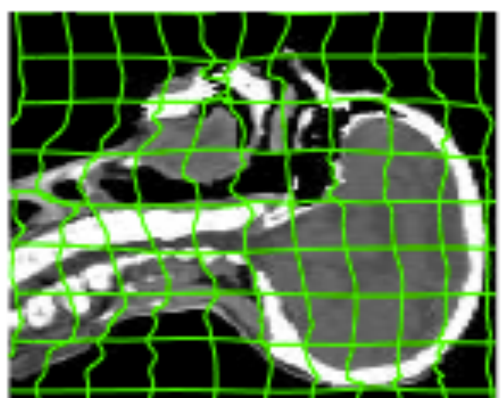




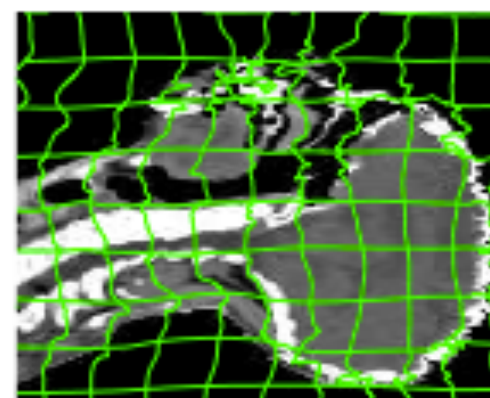
Large λ



Medium λ

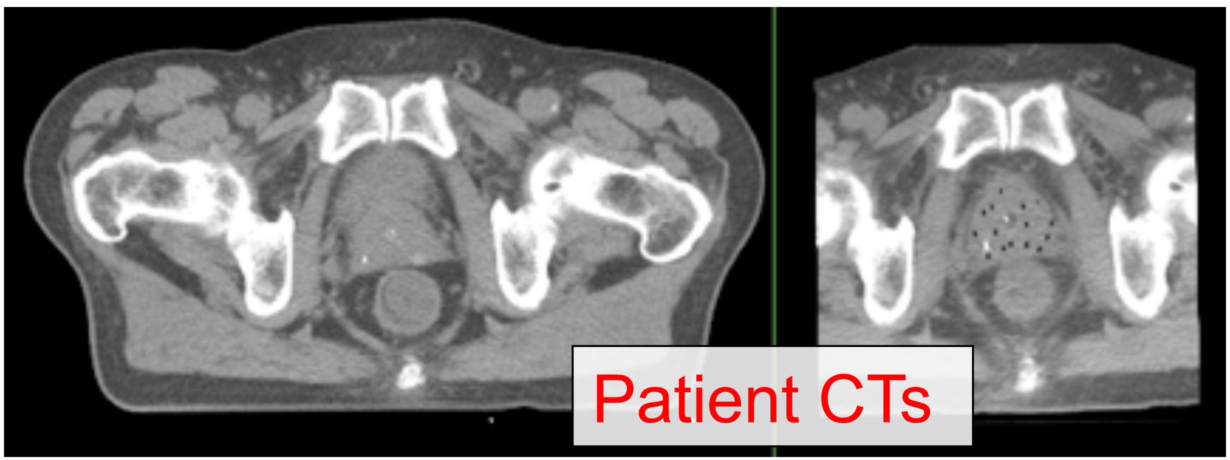


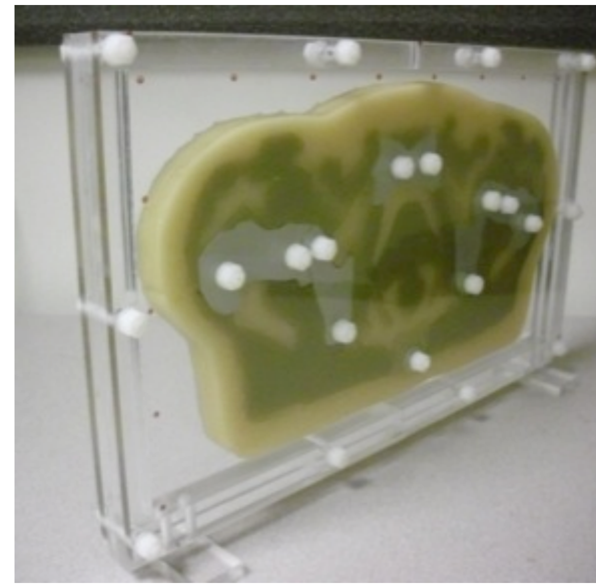
Small λ

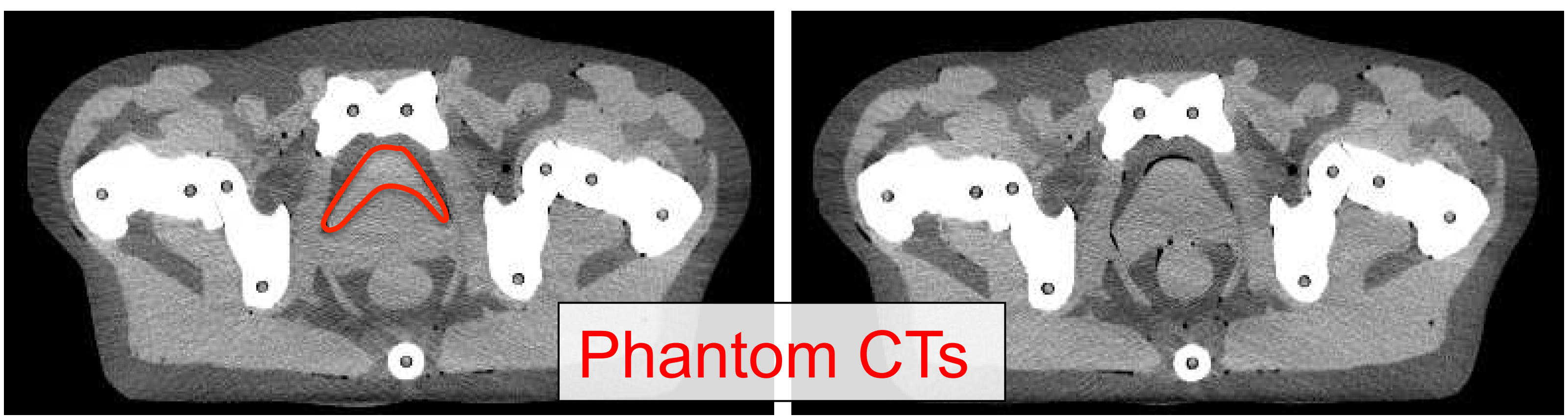


UCSF Deformable Pelvic Phantom

- Pelvic phantom to represent deformations from bladder filling.
- Realistic soft-tissue variations.

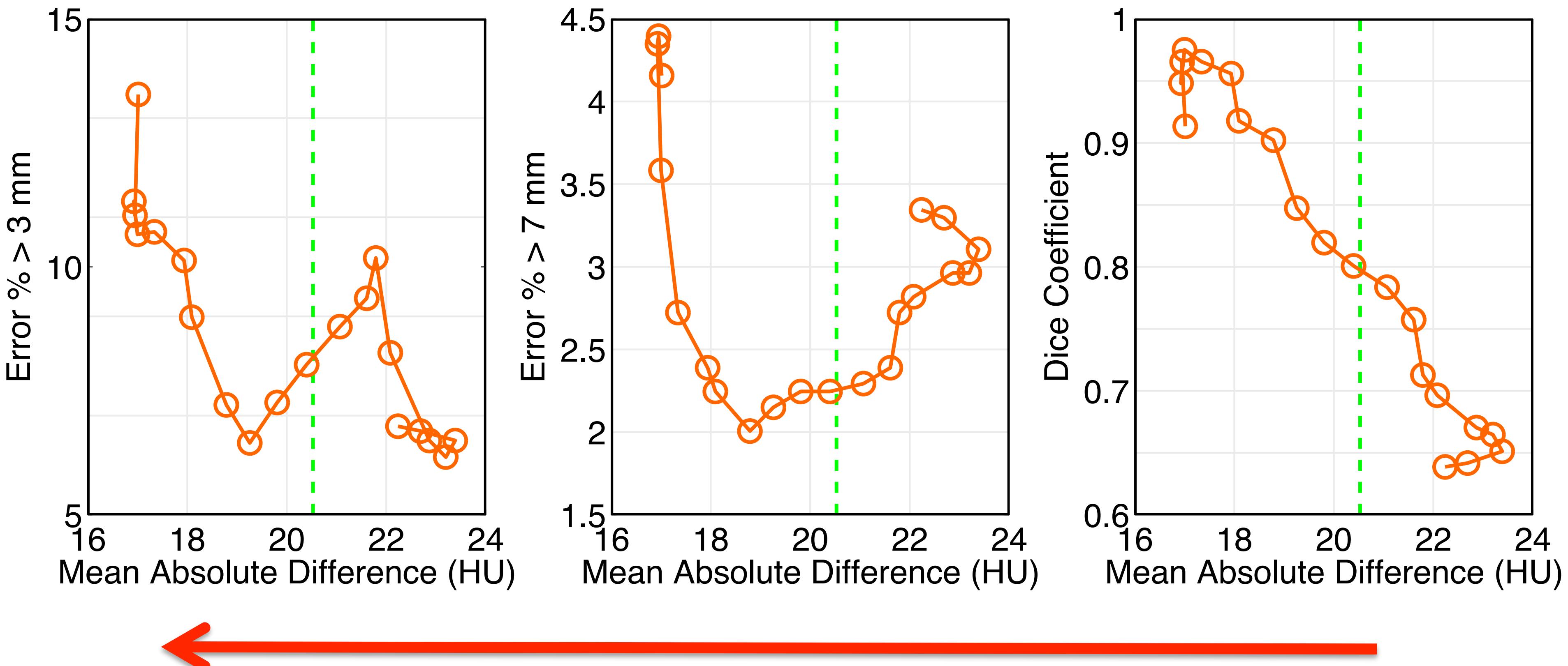






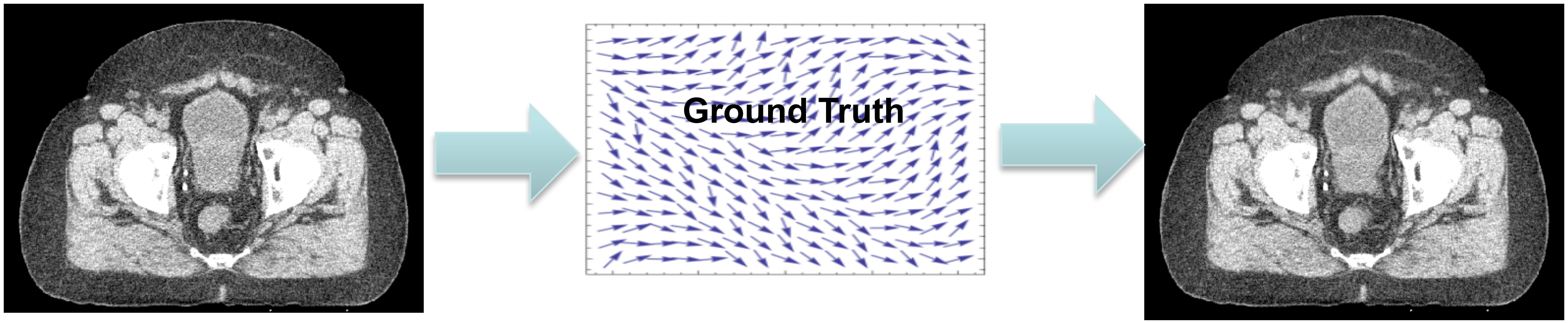
Variable Regularization Factor

- MIM tested with a variable regularization factor (20 values).
- Overall spatial error (relevant for dose mapping) optimized with a different factor than contour transfer.



More Image Similarity

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.

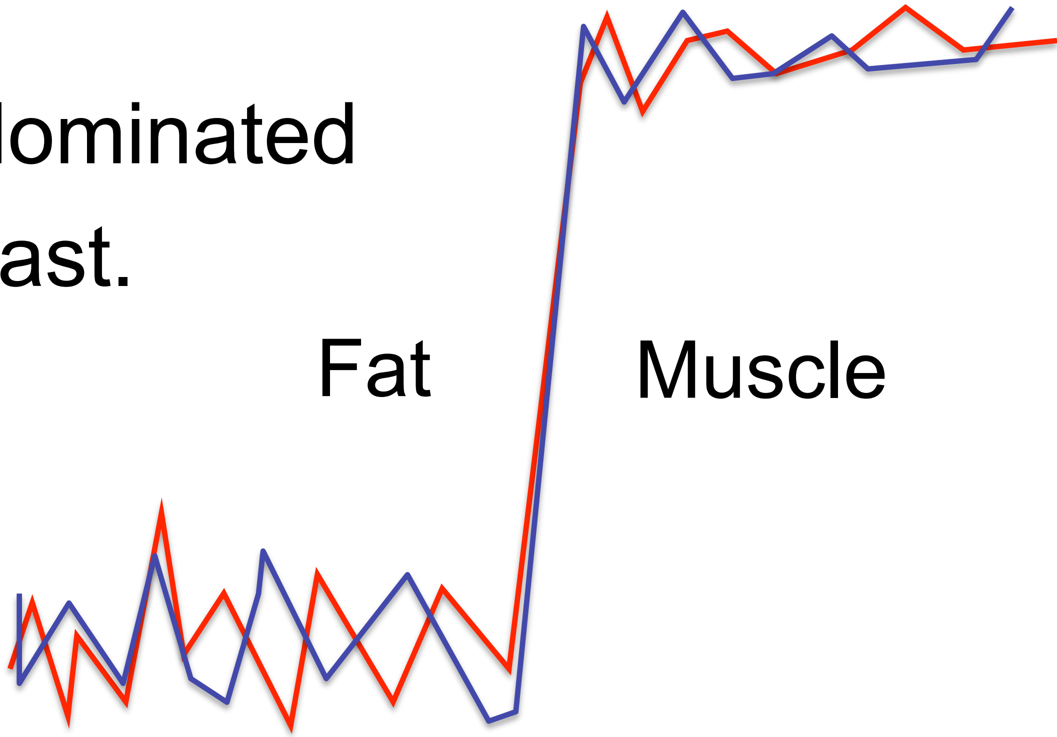
2. Landmark Tracking. $\neq$ Actual Accuracy.

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. ≠ Actual Accuracy.

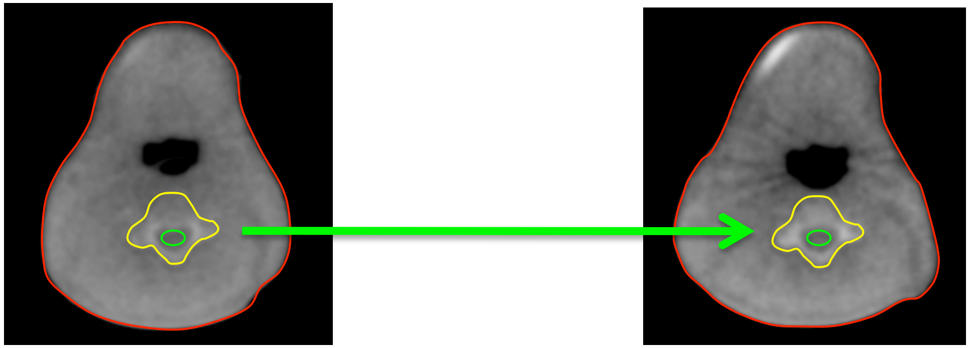
3. Image similarity.

Available for Patients

DIR QA for a patient.

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

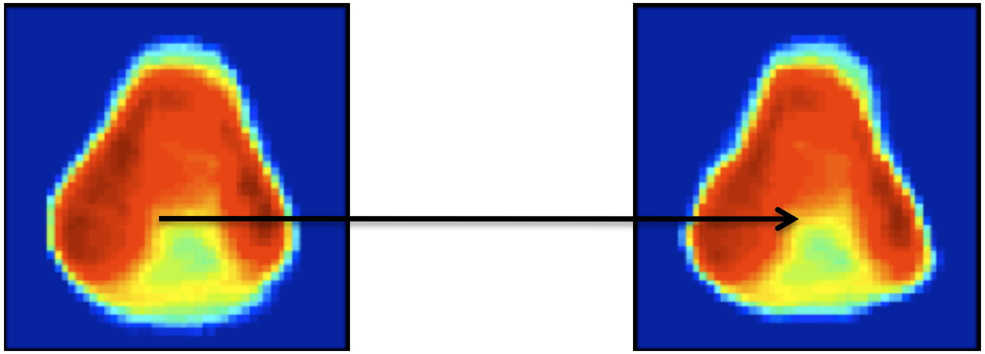
- For contour transfer, you simply need to visually inspect these contours.



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.



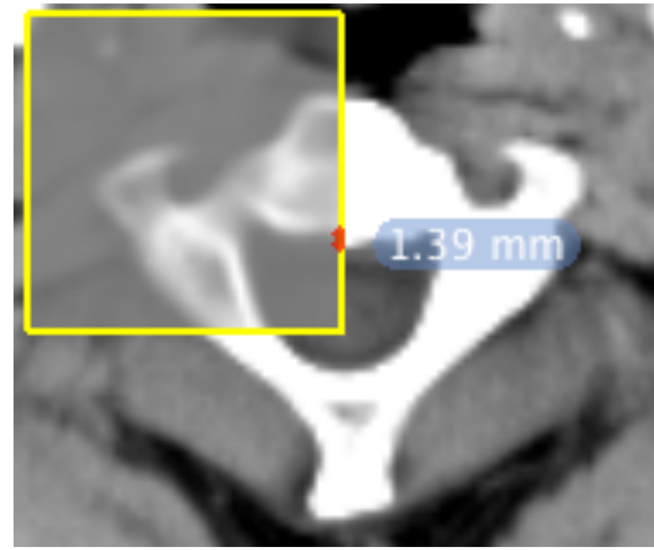
DIR QA for a patient.

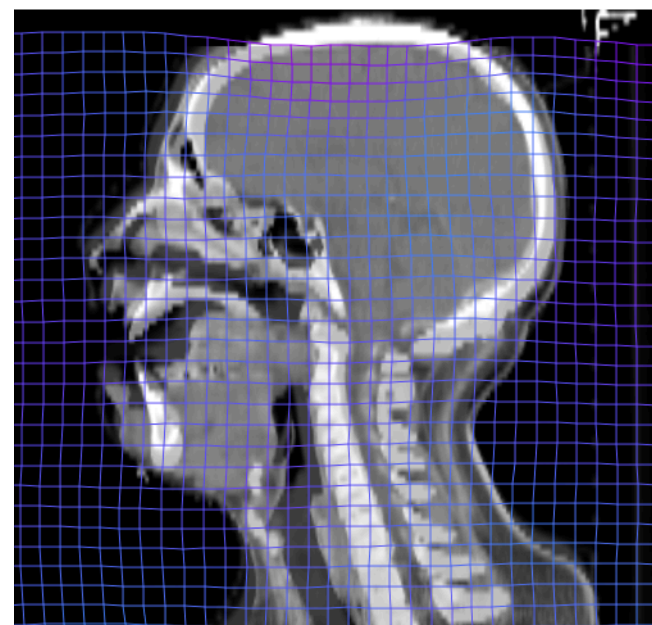
•Evaluating DIR spatial accuracy must be broken into two parts:

•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)





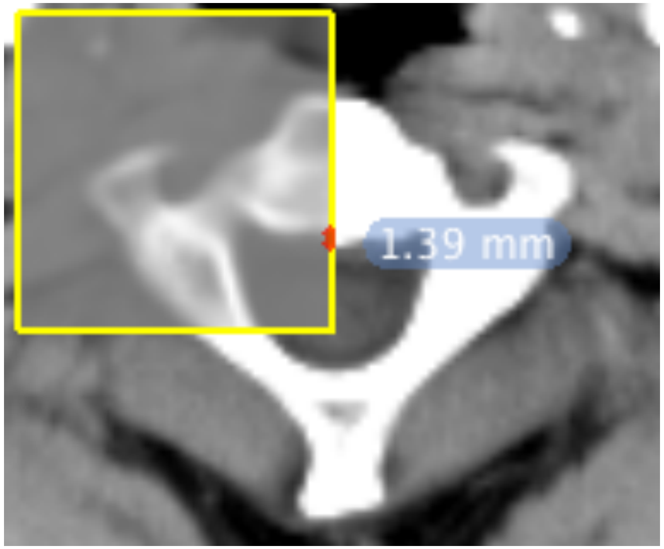
Quantitative DIR QA for a patient.

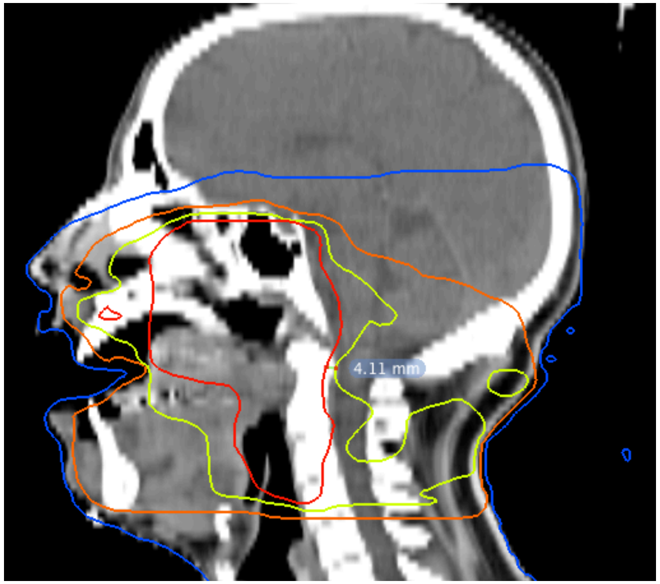
- You can estimate the spatial accuracy from these two terms and multiply by local dose gradient.

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)

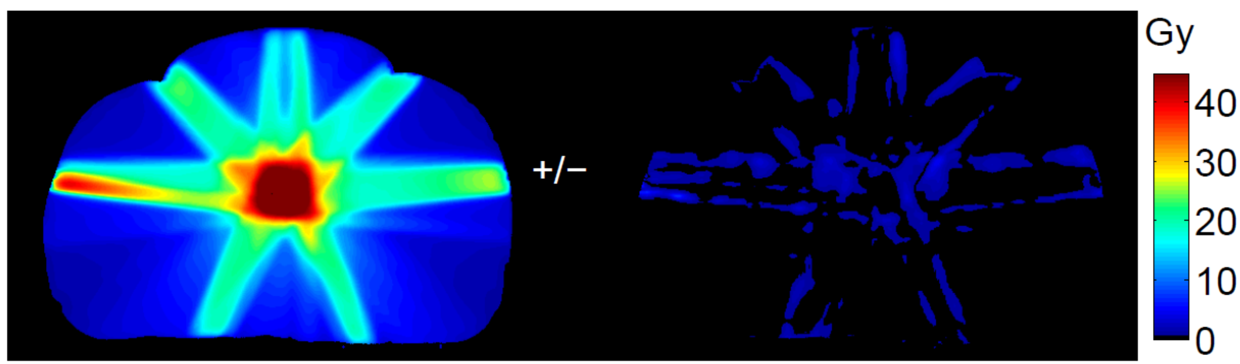




Automated DIR QA for a patient.

- 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.

- **H. Kim** et al. “Validating Dose Uncertainty Estimates Produced by AUTODIRECT, An Automated Program to Evaluate Deformable Image Registration Accuracy”, **SU-E-J-92**, Exhibit Hall

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

20%	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.

10

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.