Quality control in MR-only radiotherapy treatment planning using multi-task learning and uncertainty estimation

Felix Bragman[1,2], Ph.D.

1. Centre for Medical Image Computing, University College London
2. Artificial Medical Intelligence Group, School of Biomedical Engineering and Imaging Sciences, King's College London

29/03/2019

MR in Radiotherapy
British Institute of Radiology

Talk summary

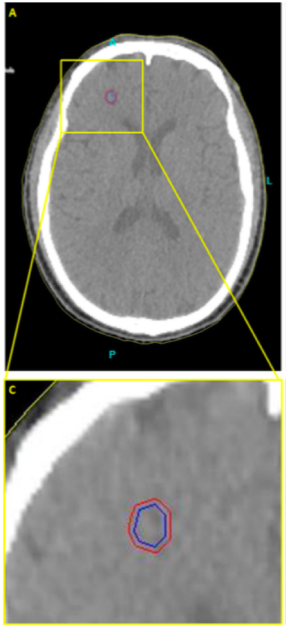
- MR-only radiotherapy treatment planning
- Methods for synthetic CT generation and organ at risk segmentation
- Deep learning for MR-only radiotherapy treatment planning [MICCAI & MIDL 2018]

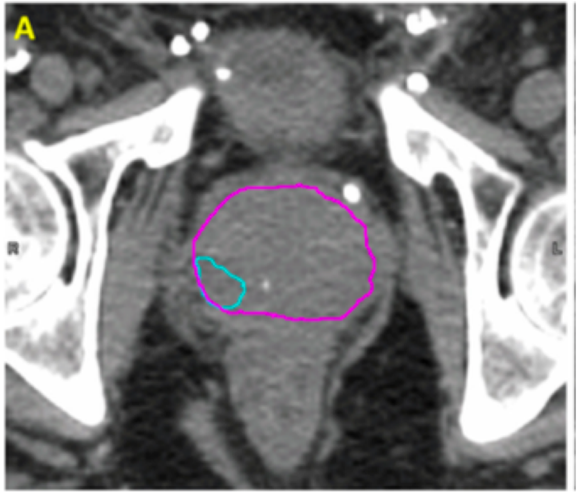
Automated CT synthesis and OAR segmentation from MRI

- Treatment planning requires both computed tomography (CT) and magnetic resonance imaging (MRI)

With only CT

- No contrast between tumour and surrounding normal tissue
- Errors in the delineation of OARs

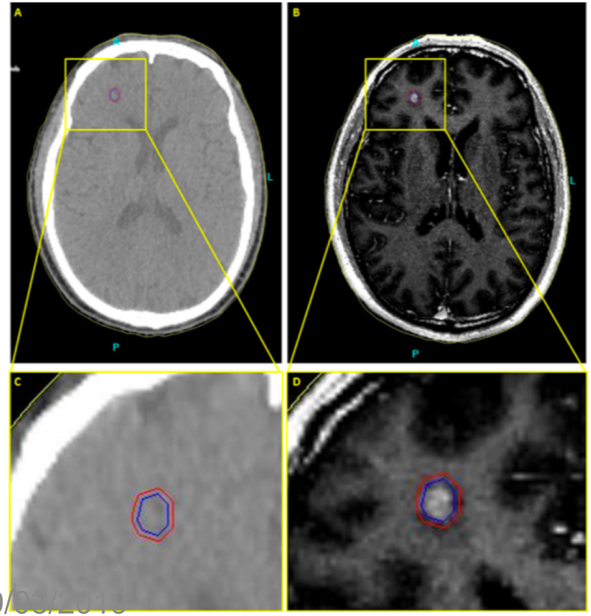


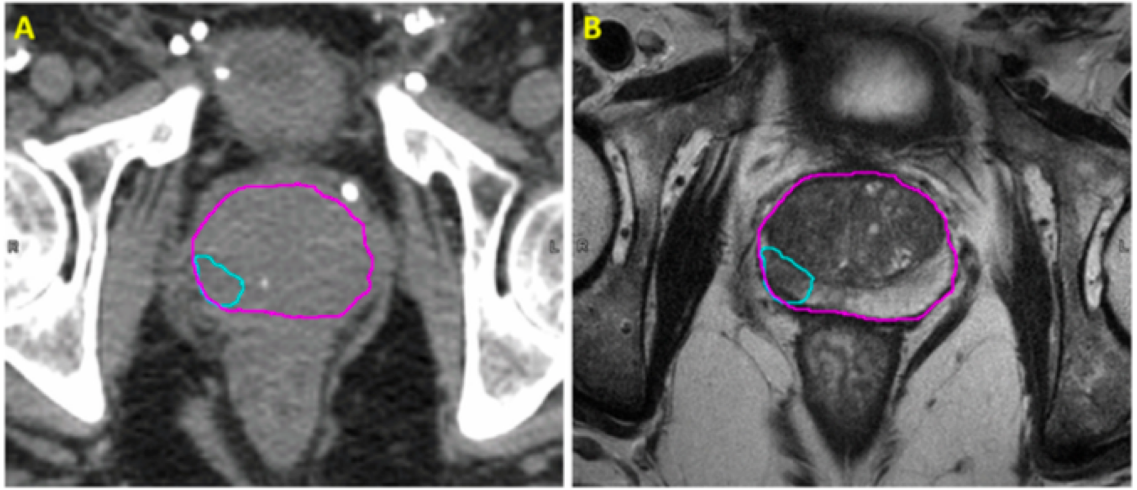


Figures from Phys.Med.Bio 63(2018) 05TR01

Automated CT synthesis and OAR segmentation from MRI

With MRI





MR in Radiotherapy
British Institute of Radiology

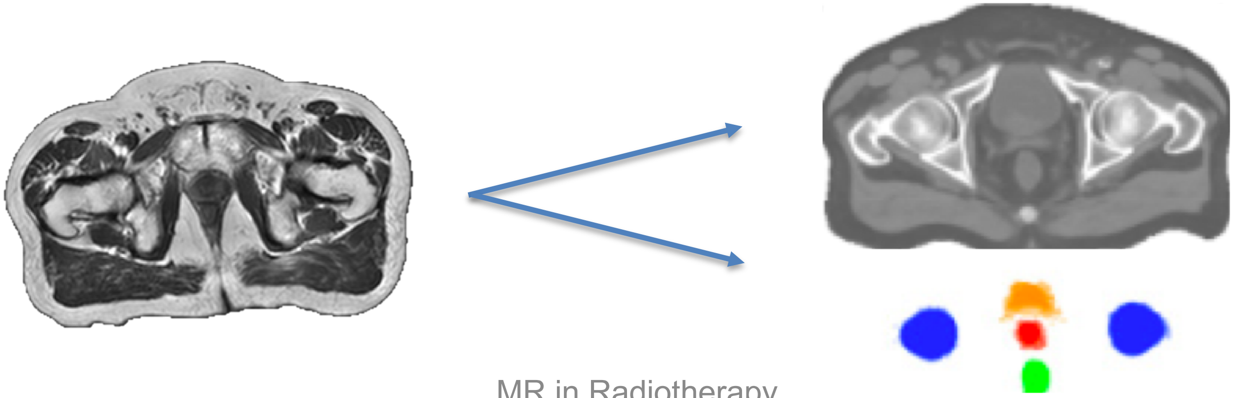
Figures from Phys.Med.Bio 63(2018) 05TR01

Difficulties introduced with MR-CT registration

- Image registration of MR to CT scans
  - Introduces geometrical uncertainty: ~2mm in brain and ~2-3mm for prostate [Haider et al. 2018]
  - Systematic errors → shift high dose regions away from target + geometric miss
- Unnecessary CT scanning → radiation dose, patient time and imaging costs

MR-only radiotherapy treatment planning

- No image registration – generate a synthetic CT scan directly from MR
- Segmentation of key structures performed using MR scans with high soft-tissue contrast

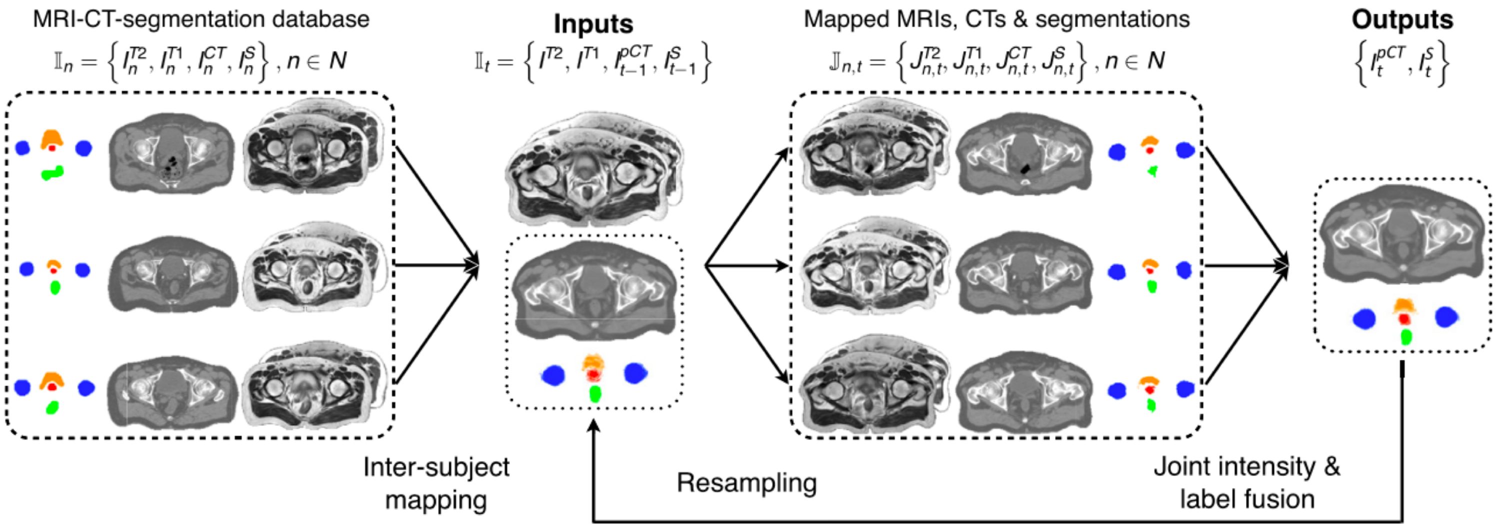


MR-only radiotherapy treatment planning

- Traditional methods
  - Label fusion: registration, propagation and fusion [Burgos et al., 2017]
- Machine learning
  - Generative models [Cardoso et al., 2015]
  - Random forest regression [Jog et al., 2017]
  - Convolutional neural networks [Wolterink et al., 2017]

Traditional methods for synCT generation and OAR segmentation

- Registration, propagation and fusion [Burgos et al., 2017]



Traditional methods for synCT generation and OAR segmentation

- Limitations
  - Data sharing: the algorithm by itself is useless..
  - No concept of uncertainty: what are the errors in the synCT and segmentation?
  - Fully deterministic
  - Requires inter-patient registration (>10) at every iteration of the algorithm

Deep learning for synCT generation and OAR segmentation

Limitations of old methods

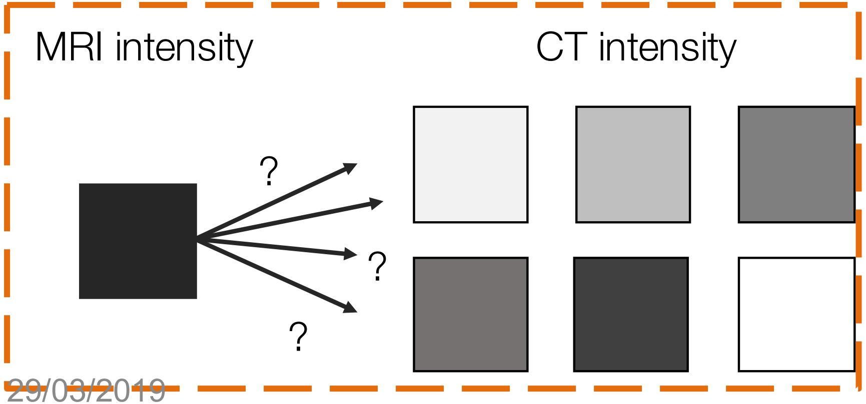
1. Data sharing issue
2. Fully deterministic system
3. No concept of uncertainty
4. Inter-patient registration is required

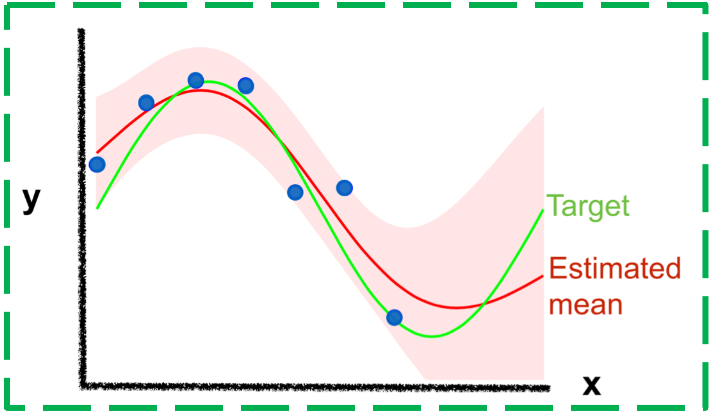
Using deep learning

1. Share derived model-parameters
2. Fully probabilistic – knowledge of the model that generates the synCT or segmentation
3. Model uncertainty in the process
4. Very fast! (for 1 patient: 10 seconds versus 24hours)

Our work

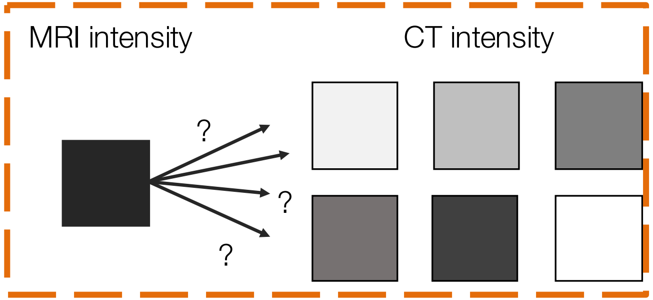
- Desirable properties of the CNN
  a) Accurate prediction for the synCT and the OAR segmentations
  b) Knowledge of the uncertainty in the predictions to be exploited for quality control
  c) Ability to sample from the model to generate realistic predictions for probability dose delivery

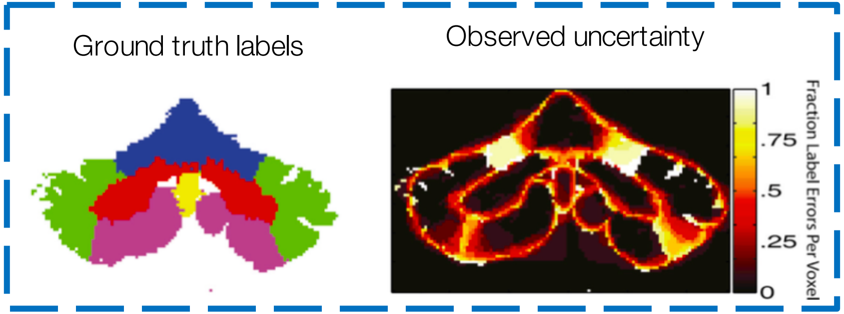




What is task uncertainty?

- Inherent ambiguity in the problem
- Uncertainty is spatial varying e.g. organ segmentation
- We want to be able to predict this uncertainty
- Why?
  - It can improve the quality of the predictions
  - Knowledge of this uncertainty can be used for quality control in the synthetic CT

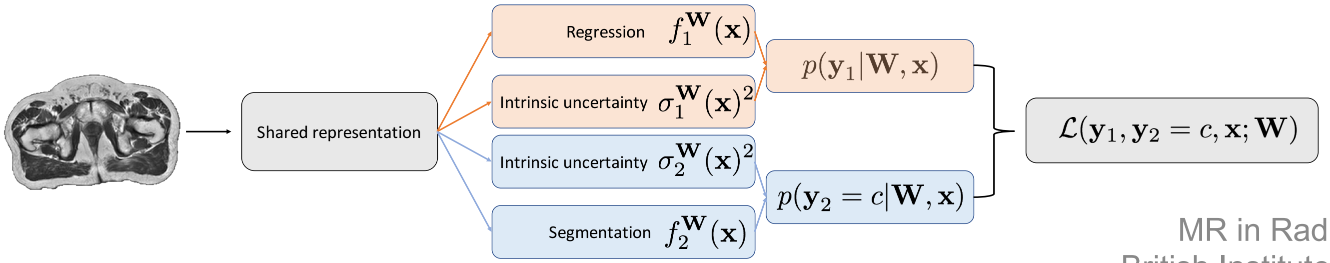




Our contribution

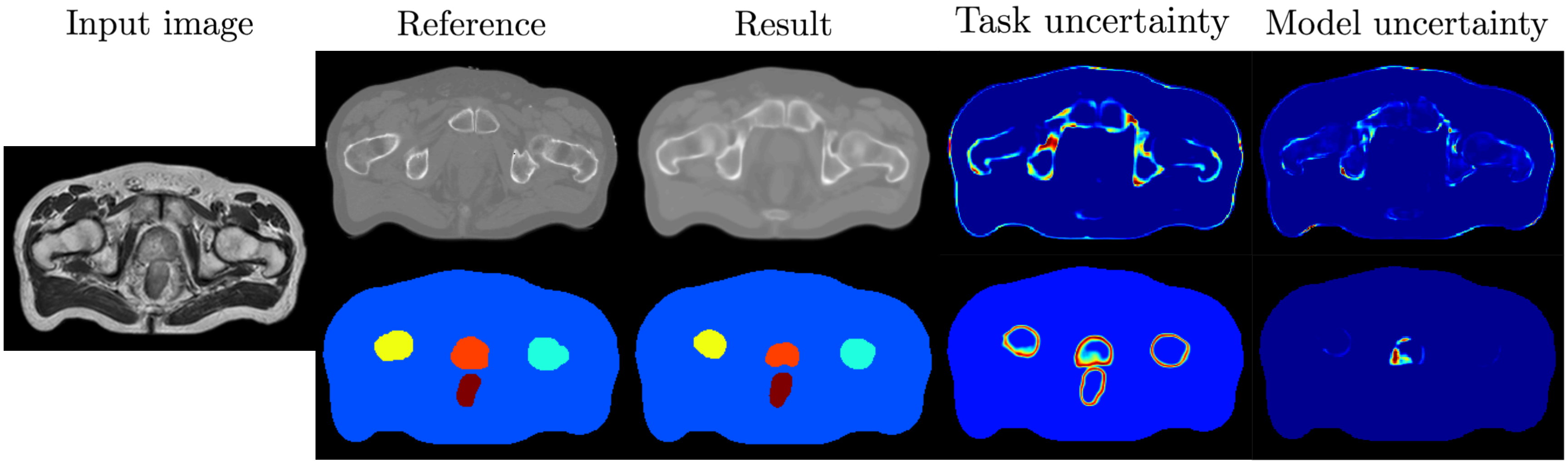
- Probabilistic multi-task network
  - Shared network + regression and segmentation specific branches
- Predict task-specific uncertainty for regression and segmentation to analyse model predictions
- Applied Bayesian modelling to enable stochastic sampling at test time

$$\mathcal{L}(\mathbf{y}_1, \mathbf{y}_2 = c, \mathbf{x}; \mathbf{W}) = \frac{||\mathbf{y}_1 - f_1^{\mathbf{W}}(\mathbf{x})||^2}{2\sigma_1^{\mathbf{W}}(\mathbf{x})^2} + \frac{\mathrm{CE}(f_2^{\mathbf{W}}(\mathbf{x}), \mathbf{y}_2 = c)}{2\sigma_2^{\mathbf{W}}(\mathbf{x})^2} + \log\left(\sigma_1^{\mathbf{W}}(\mathbf{x})^2 \sigma_2^{\mathbf{W}}(\mathbf{x})^2\right)$$



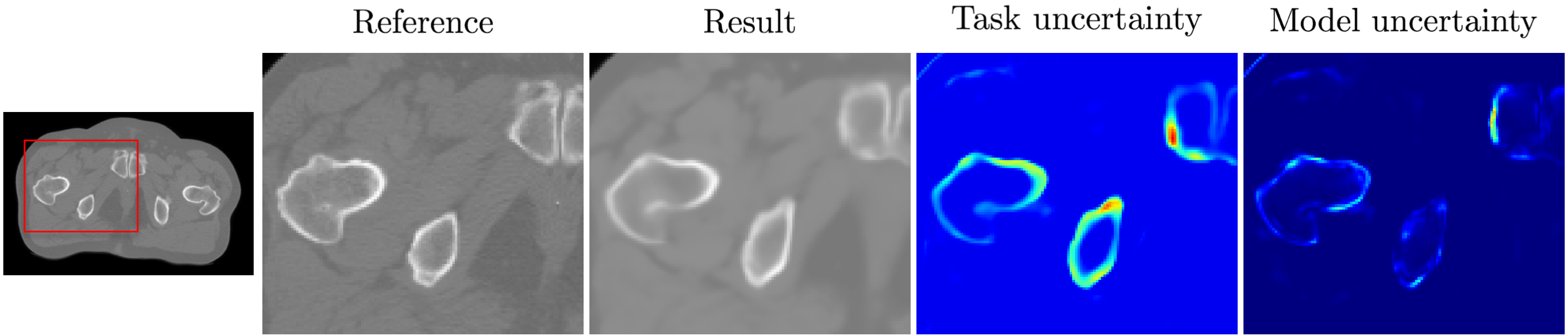
Experiment on 15 prostate cancer patients

- 3-fold cross-validation for training and testing



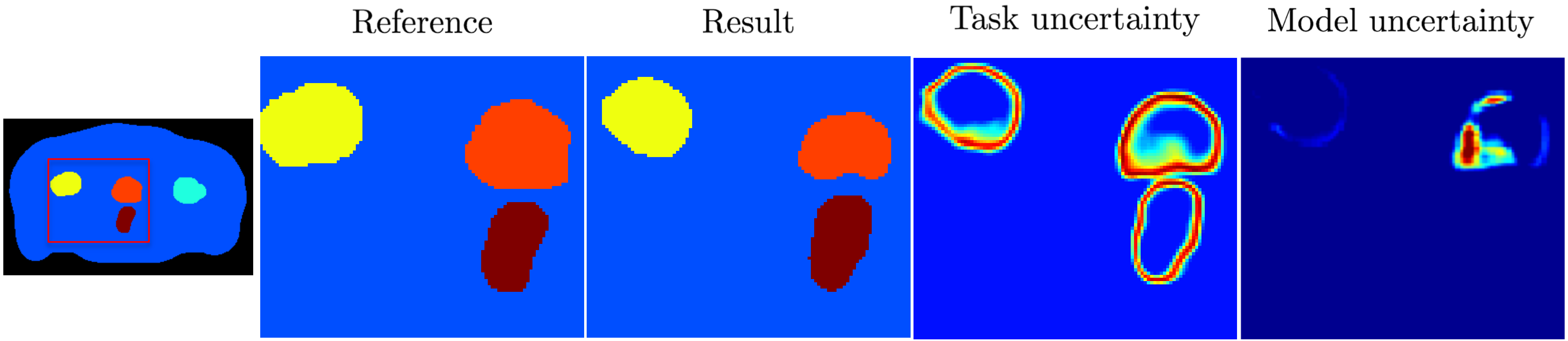
Experiment on 15 prostate cancer patients

- 3-fold cross-validation for training and testing



Experiment on 15 prostate cancer patients

- 3-fold cross-validation for training and testing



Main results

1. Our model outperforms all baseline models including label fusion [Burgos et al., 2017]

Models	All	Bone	L femur	R femur	Prostate	Rectum	Bladder
			Regression - synCT - Mean Absolute Error (HU)				
HighResNet [7]	**48.1(4.2)**	**131(14.0)**	78.6(19.2)	**80.1(19.6)**	**37.1(10.4)**	63.3(47.3)	**24.3(5.2)**
HighResNet + dropout	**47.4(3.0)**	**130(12.1)**	78.0(14.8)	77.0(13.0)	**36.5(7.8)**	67(44.6)	**24.1(7.5)**
HighResNet + dropout + hetero [6]	44.5(3.6)	128(17.1)	75.8(20.1)	74.2(17.4)	31.2(7.0)	56.1(45.5)	17.8(4.7)
Multi-task + homo noise weighting [1]	44.3(3.1)	126(14.4)	74.0(19.5)	73.7(17.1)	29.4(4.7)	58.4(48.0)	18.2(3.5)
Multi-atlas propagation [5]	45.7(4.6)	125(10.3)	-	-	-	-	-
Multi-task + dropout + hetero	43.3(2.9)	121(12.6)	69.7(13.7)	67.8(13.2)	28.9(2.9)	55.1(48.1)	18.3(6.1)

Multi-atlas propagation [5]	45.7(4.6)	125(10.3)
Multi-task + dropout + hetero	43.3(2.9)	121(12.6)

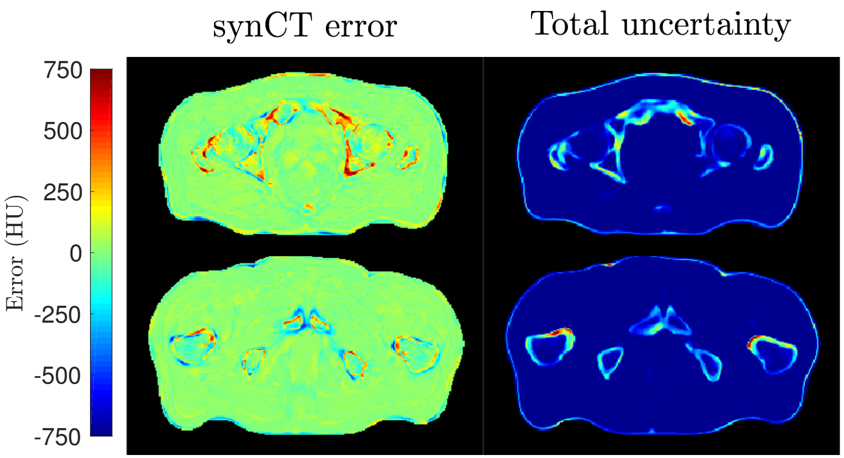
Main results

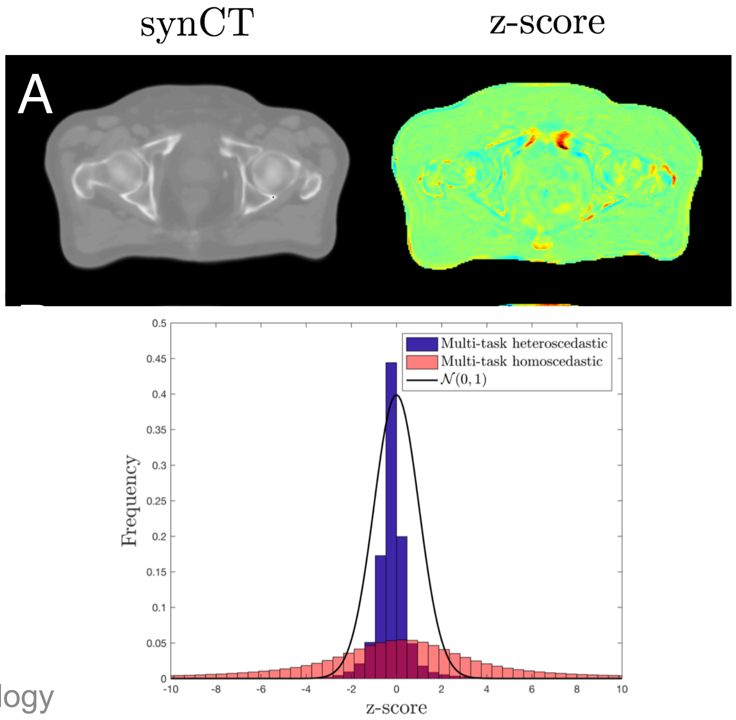
- Equivalent results with state of the art in segmentation
- Label fusion method used 3D T1/T2 scans…we trained only using 2D slices from T2

	L femur	R femur	Prostate	Rectum	Bladder
Multi-atlas propagation [5]	0.89(0.02)	0.90(0.01)	0.73(0.06)	0.77(0.06)	0.90(0.03)
Multi-task + dropout + hetero	0.91(0.02)	0.91(0.02)	0.70(0.06)	0.74(0.12)	0.93(0.04)

Using uncertainty for quality control

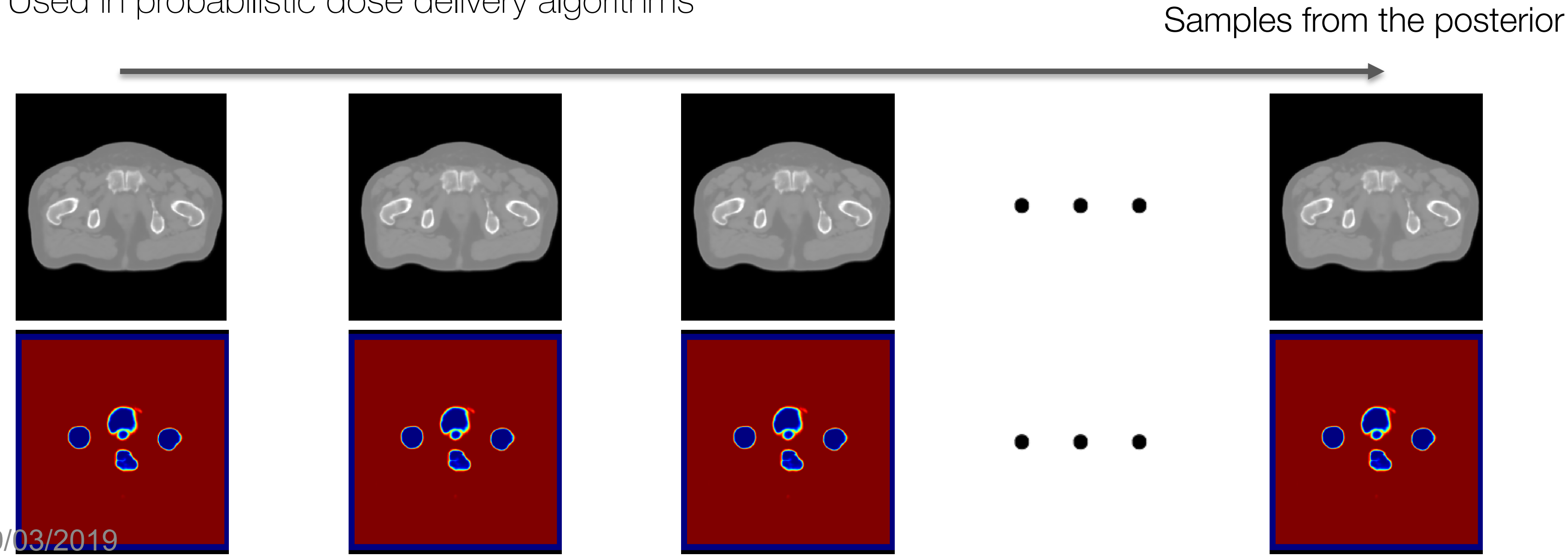
- Predicted uncertainty in the synCT correlates strongly with areas of high error
- Uncertainty is well calibrated





Sampling from the model for probabilistic dose delivery estimations

- Bayesian model so we can sample from the posterior distribution at test time
- Generate multiple realistic realisations of the synCT given an MR scan
- Used in probabilistic dose delivery algorithms



Thank you for listening

- [1] Bragman et al. *Uncertainty in multitask learning: joint representations for probabilistic MR-only radiotherapy planning*, MICCAI 2018
- [2] Bragman et al. Quality control in radiotherapy-treatment planning using multi-task learning and uncertainty estimation, MIDL 2018

- Code will be released open-source as part of the NiftyNet package (www.niftynet.io)
- Download @ **`pip install niftynet`**