The VAMPIRE Challenge:

Preliminary results from a multi-institutional study of CT ventilation image accuracy

John Kipritidis, CINSW Early Career Fellow
Radiation Physics Laboratory
UNIVERSITY OF SYDNEY
SYDNEY MEDICAL SCHOOL

SAM JOINT IMAGING-THERAPY SCIENTIFIC SYMPOSIUM
AAPM Annual Meeting, July 31st - August 4th 2016
Washington DC, USA

Learning Objectives

① To understand the need for common datasets to validate and improve CT ventilation imaging (CTVI)

② To learn about common tools for evaluating CTVI accuracy

③ To learn about a global initiative to identify the most accurate CTVI algorithms

2

Outline

› Goals and barriers for cross-modality validation of CTVI

› *VAMPIRE:* Realizing a multi-institutional validation dataset for CTVI

› *VAMPIRE Challenge:* Preliminary results based on USYD algorithm

3

SYDNEY

Outline

› Goals and barriers for cross-modality validation of CTVI

› *VAMPIRE:* Realizing a multi-institutional validation dataset for CTVI

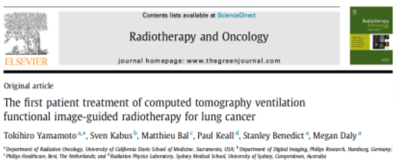
› *VAMPIRE Challenge:* Preliminary results based on USYD algorithm

4

SYDNEY

Goals and barriers for cross-modality validation

- CTVI-guided *functional avoidance lung cancer RT* is now a clinical reality and can have a material impact on functional dosimetry.


Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

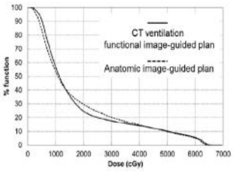
ELSEVIER

Original article

The first patient treatment of computed tomography ventilation functional image-guided radiotherapy for lung cancer

Tokihiro Yamamoto[a,*], Sven Kabus[b], Matthieu Bal[c], Paul Keall[d], Stanley Benedict[a], Megan Daly[a]

[a] Department of Radiation Oncology, University of California Davis School of Medicine, Sacramento, USA; [b] Department of Digital Imaging, Philips Research, Hamburg, Germany; [c] Philips Healthcare, Best, The Netherlands; and [d] Radiation Physics Laboratory, Sydney Medical School, University of Sydney, Camperdown, Australia



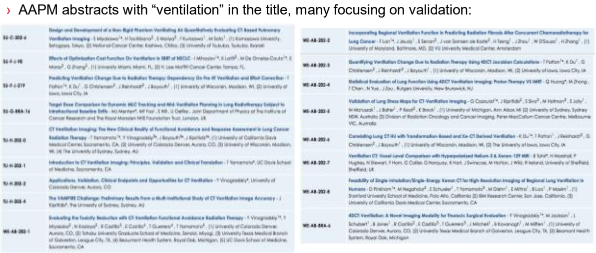
T. Yamamoto *et al.* (2016)
Radiother Oncol

5

SYDNEY

Goals and barriers for cross-modality validation

› AAPM abstracts with “ventilation” in the title, many focusing on validation:



6

Goals and barriers for cross-modality validation

- AAPM abstracts with "ventilation" in the title, many focusing on validation:

Goals and barriers for cross-modality validation

- Variable results across different clinical / pre-clinical validation studies:

Study	Modality	Subjects	Key finding
Yamamoto *et al.* [1]	PFT	15 human subjects	Moderate correlations (range 0.43-0.73) between CT ventilation-defined

SAM hint: Are these studies directly comparable?

- Different *subjects*,
- Different *4DCT protocols*,
- Different *validation metrics*,
- Different *"ground truth" ventilation modalities*.

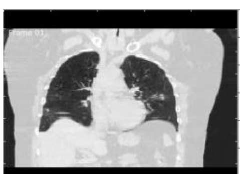
Abbreviations: PFT = pulmonary function test, ^{99m}Tc-DTPA = technetium-99m-labeled diethylenetriamine pentaacetate, DSC = Dice similarity coefficient.

S. Kida *et al.* (2016) Radiother Oncol 118 (3) pg. 521-527

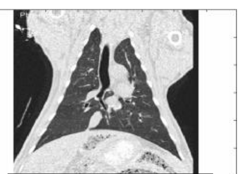
Goals and barriers for cross-modality validation

- 90% of clinical 4DCT scans suffer image artifacts due to irregular breathing.

Yamamoto et al. (2008) IJROBP 72 (4) 1250-1258



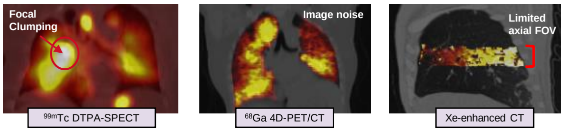
Free-breathing 4DCT



Mechanically ventilated sheep

Goals and barriers for cross-modality validation

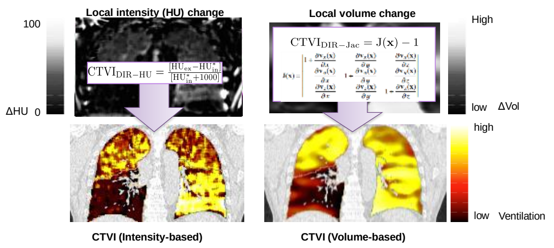
› Different clinical ventilation images have different pros/cons in terms of image quality:



10

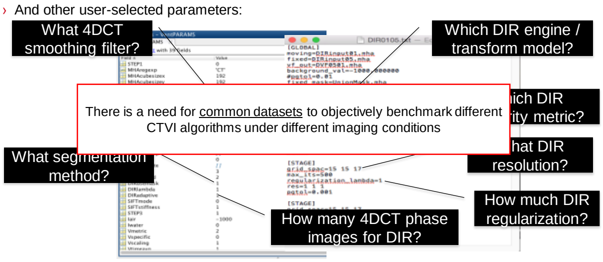
Goals and barriers for cross-modality validation

› There are multiple methods to quantify breathing-induced air volume change using DIR:



Goals and barriers for cross-modality validation

› And other user-selected parameters:



Outline

› Goals and barriers for cross-modality validation of CTVI

› *VAMPIRE:* Realizing a multi-institutional validation dataset for CTVI

› *VAMPIRE Challenge:* Preliminary results based on USYD algorithm

13

Outline

› Goals and barriers for cross-modality validation of CTVI

› *VAMPIRE:* Realizing a multi-institutional validation dataset for CTVI

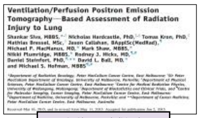
› *VAMPIRE Challenge:* Preliminary results based on USYD algorithm

14

VAMPIRE: Realizing a common validation dataset

- ***"Ventilation Archive for Medical Pulmonary Image Registration Evaluation"***
  50 pairs of 4DCT and 'ground truth' ventilation scans from three institutions (more soon)

Peter Mac


Ventilation/Perfusion Positron Emission Tomography—Based Assessment of Radiation Injury to Lung

S. Siva *et al. (*2015)
IJROBP 93 (2) pg. 408-17

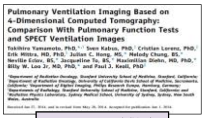
The University of Iowa


Medical Image Analysis

Registration-based estimates of local lung tissue expansion compared to xenon CT measures of specific ventilation

J. Reinhardt *et al. (*2008)
Med. Imag. Analysis 12 (6) pg. 752-63

Stanford School of Medicine

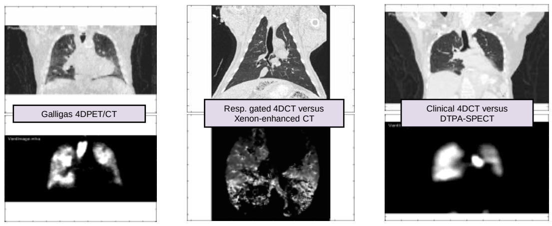

Pulmonary Ventilation Imaging Based on 4-Dimensional Computed Tomography: Comparison With Pulmonary Function Tests and SPECT Ventilation Images

T. Yamamoto *et al. (*2014)
IJROBP 90 (2) pg. 414-22

THE UNIVERSITY OF SYDNEY

VAMPIRE: Realizing a common validation dataset

- VAMPIRE includes 4DCT / ground truth imaging acquired under different conditions:



THE UNIVERSITY OF SYDNEY

VAMPIRE: Realizing a common validation dataset

Image quality++ (→)

Image quality++ (↓)

CT/GT Modalities	DTPA V-SPECT	Technegas V-SPECT	Galligas 4DPET/CT	Xenon CT	MAA Q-SPECT	Hyp. Gas MRI
4D cone-beam CT	-	-	-	-	-	-
Low-dose 4DCT	-	-	25	-	-	-
Treatment planning 4DCT	21	-	-	-	-	-
Regularity gated 4DCT	-	-	-	4	-	-
Breath hold CT	-	-	-	-	-	-

LEGEND	Inaugural VAMPIRE Challenge!	N/A

THE UNIVERSITY OF SYDNEY

VAMPIRE: Realizing a common validation dataset

Image quality++ (→)

Image quality++ (↓)

CT/GT Modalities	DTPA V-SPECT	Technegas V-SPECT	Galligas 4DPET/CT	Xenon CT	MAA Q-SPECT	Hyp. Gas MRI
4D cone-beam CT	-	11	-	-	11	-
Low-dose 4DCT	-	-	25	-	-	-
Treatment planning 4DCT	21	11	18	-	32	-
Regularity gated 4DCT	-	-	-	4	-	-
Breath hold CT	-	-	18	-	-	Up to 70

LEGEND	Inaugural VAMPIRE Challenge!	VAMPIRE II (Date TBA)	N/A

Outline

› Goals and barriers for cross-modality validation of CTVI

› *VAMPIRE:* Realizing a multi-institutional validation dataset for CTVI

› *VAMPIRE Challenge:* Preliminary results based on USYD algorithm

19

Outline

› Goals and barriers for cross-modality validation of CTVI

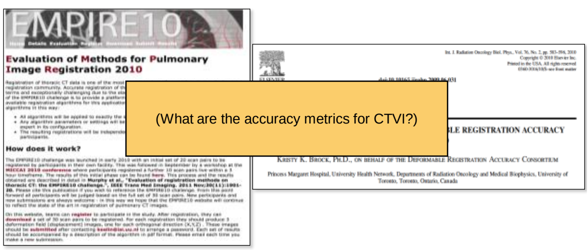
› *VAMPIRE:* Realizing a multi-institutional validation dataset for CTVI

› *VAMPIRE Challenge:* Preliminary results based on USYD algorithm

20

Inaugural *VAMPIRE Challenge*

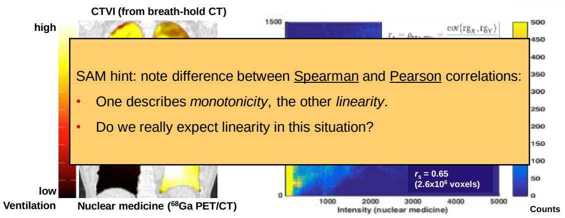
- The VAMPIRE Challenge is inspired by MIDRAS & EMPIRE10 (http://empire10.isi.uu.nl)



(What are the accuracy metrics for CTVI?)

Inaugural *VAMPIRE Challenge*

› We want high **Spearman correlation** ("r_S" or "ρ") with clinical ventilation imaging.

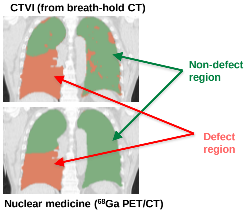


SAM hint: note difference between Spearman and Pearson correlations:

- One describes *monotonicity*, the other *linearity*.
- Do we really expect linearity in this situation?

Inaugural *VAMPIRE Challenge*

› We also want overlap of defect regions: **Dice similarity coefficient (DSC)**



DSC (i.e. fractional volume overlap) can be calculated for defect and non-defect regions:

$$\mathrm{DSC}_{\mathrm{defect}} = 2 \times \frac{|\mathrm{CTVI}_{\mathrm{defect}} \cap \mathrm{Truth}_{\mathrm{defect}}|}{|\mathrm{CTVI}_{\mathrm{defect}}| + |\mathrm{Truth}_{\mathrm{defect}}|}$$

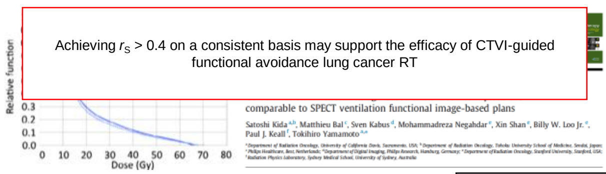
For this example:

- $\mathrm{DSC}_{\mathrm{defect}}$ = 0.67; a value *>0.5 is "moderate"*
- $\mathrm{DSC}_{\mathrm{nondefect}}$ = 0.83; a value *>0.8 is "strong"*

Inaugural *VAMPIRE Challenge*

What is a "good" correlation?

- Kida *et al.* demonstrated that r_S ~0.4 between CTVI and SPECT can produce comparable functionally adaptive treatment plans:



Achieving r_S > 0.4 on a consistent basis may support the efficacy of CTVI-guided functional avoidance lung cancer RT

S. Kida *et al.* (2016) Radiother Oncol 118 (3) pg. 521-527

Inaugural *VAMPIRE Challenge*

What is a “good” correlation?

› We found r_S = (0.66 ± 0.19) when comparing V/Q-SPECT images for 11 patients with matched ventilation/perfusion.

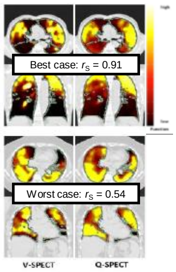


Image courtesy Dr Fiona Hegi-Johnson (Nepean Hospital, NSW Australia)

25

Inaugural *VAMPIRE Challenge*

What is the structure of the VAMPIRE Challenge? (And how to take part)

① **Download the VAMPIRE Challenge Dataset**
- Email john.kipritidis@sydney.edu.au to register.
- Receive link to Dicom or MetaImage (MHA) download.
- Dataset divided into “*Training*” / “*Validation*” components (30% / 70% split). *(Ground truth ventilation provided for Training component only)*

② **Upload your CTVI results:**
- DIR motion fields
- Processed CTVIs
- Answer a multiple-choice survey to characterize your algorithm
- Deadline is October, 2016.

③ **Voxel-based correlations performed at University of Sydney (+ others):**
- Focus on Spearman r_S and DSC values for defect/non-defect regions
- Determine which CTVI algorithm(s) achieve the best spatial accuracy

Inaugural *VAMPIRE Challenge*

○ 12 participating groups (so far!) using various commercial / open-source DIR software:

mim SOFTWARE

VARIAN medical systems | Velocity

Plastimatch

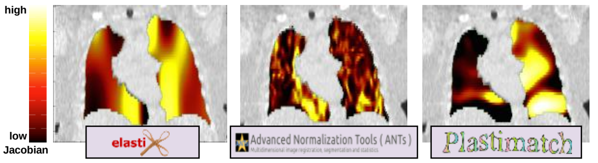
elastix

Advanced Normalization Tools (ANTs)
Multidimensional image registration, segmentation and statistics

27

Inaugural *VAMPIRE Challenge*

- Different DIR software can produce quite different lung deformation (Jacobian) maps:



VESPIR (*VEntilation via Scripted Pulmonary Image Registration*)

Sub-function	Options / features
(1) File import	○ Supports Dicom / MHA files
(2) Image cropping*,†	○ GUI-based ROI selection
(3) Gaussian filtering*	○ FWHM selection
(4) Lung segmentation*	○ Selection of intensity levels for air/water/lung † ○ Peak-exhale/inhale selection † ○ Allow mask import from file
(5) Deformable Image Registration (DIR)*	○ Supports B-spline DIR using Plastimatch or Elastix ○ Supports two DIR schemes: Inhale/Exhale and '4D' ○ Supports MSE/MI-based DIR both with/without lung masks ○ Adjustable DIR regularization ○ 1-6 DIR resolution stages † ○ Supports user parameter files
(6) Ventilation quantification	○ Supports four CTVI metrics (intensity or volume based) ○ Supports 4D time-averaging* ○ Mask-preserving median filter* ○ Avoidance map generation ○ Dicom output*
(7) Validation of DIR and/or CTVI*	○ In-built visual DIR evaluation ○ Spearman correlation of CTVI with clinical ventilation scan

Kipritidis *et al.* (2016)
2016 ISBI pg. 939-942

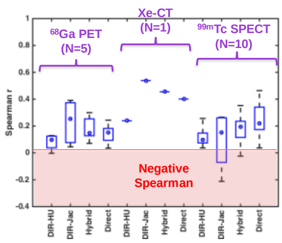
We tested a number of different CTVI calculations:

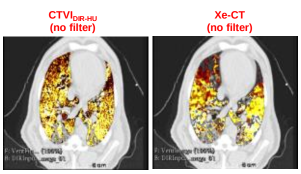
Metric	Expression
$\mathrm{CTVI}_{\mathrm{DIR-HU}}$:	$\frac{[\mathrm{HU}_{\mathrm{ex}}(\mathbf{x}) - \mathrm{HU}^{*}_{\mathrm{in}}(\mathbf{x}+\mathbf{v})]}{[\mathrm{HU}^{*}_{\mathrm{in}}(\mathbf{x}+\mathbf{v})+1000]}$
$\mathrm{CTVI}_{\mathrm{DIR-Jac}}$:	$[\mathrm{Jac}(\mathbf{x},\mathbf{v}) - 1]$
$\mathrm{CTVI}_{\mathrm{Hybrid}}$:	$[\mathrm{HU}_{\mathrm{ex}}(\mathbf{x})/\mathrm{Jac}(\mathbf{x},\mathbf{v}) - \mathrm{HU}_{\mathrm{in}}(\mathbf{x})]$
$\mathrm{CTVI}_{\mathrm{Direct}}$:	$\sum_{\phi=1}^{N} \left\lceil \frac{\mathrm{HU}_{\phi}(\mathbf{x})}{-1000} \times \frac{\mathrm{HU}_{\phi}(\mathbf{x})+1000}{1000} \right\rceil /N$

(No algorithm "tuning" per se; visual checking of DIR)

Preliminary *VAMPIRE Challenge* results (*VESPIR*)

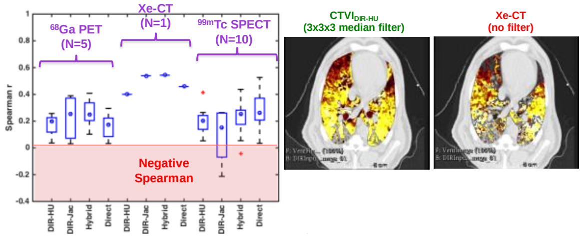
- Spearman correlations (Training data only)





Preliminary *VAMPIRE Challenge* results (*VESPIR*)

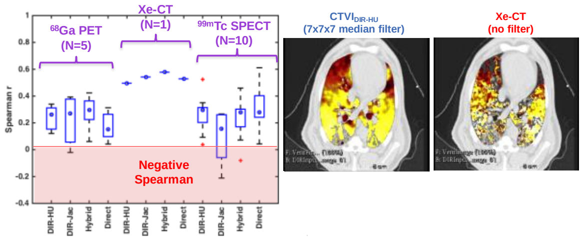
- Spearman correlations (Training data only)



31

Preliminary *VAMPIRE Challenge* results (*VESPIR*)

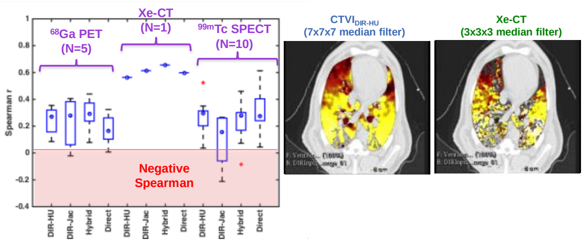
- Spearman correlations (Training data only)



32

Preliminary *VAMPIRE Challenge* results (*VESPIR*)

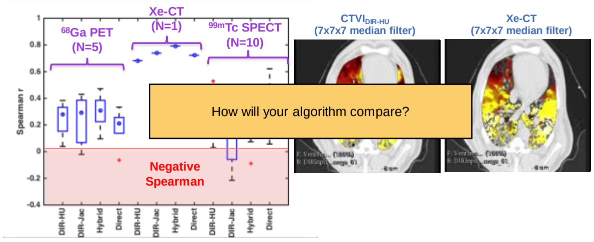
- Spearman correlations (Training data only)



33

Preliminary *VAMPIRE Challenge* results (*VESPIR*)

- Spearman correlations (Training data only)



Inaugural *VAMPIRE Challenge*

What are the benefits of participating in the *VAMPIRE Challenge*?

Benefit of providing your CTVI outputs:
- Learn how your CTVI algorithm fares against others.

Benefit of providing data:
- Learn which CTVI methods are most accurate for your imaging data.

Following each VAMPIRE Challenge:
- Where possible, full data will be publically released for purposes of CTVI validation

Take home message

› Validation remains a core focus of CTVI research.

› The VAMPIRE dataset will begin to address the need for common validation datasets.

› Please contact us to participate in the Inaugural *VAMPIRE Challenge*

→ Email john.kipritidis@sydney.edu.au to take part!

Acknowledgements

THE UNIVERSITY OF SYDNEY

- Paul J. Keall
- John Kipritidis
- Henry C. Woodruff
- Enid Eslick
- Fiona Hegi-Johnson

Peter Mac

- Michael S. Hofman
- Shankar Siva
- Jason Callahan
- Nicholas Hardcastle

THE UNIVERSITY OF IOWA

- Joseph M. Reinhardt
- Gary E. Christensen

STANFORD SCHOOL OF MEDICINE

- Bill W. Loo

UC DAVIS COMPREHENSIVE CANCER CENTER

- Tokihiro Yamamoto

The University Of Sheffield.

- Bilal Tahir
- Jim Wild
- Rob Ireland

- Kristy Brock
- Guillaume Cazoulat

Disclosure

This work has been supported by a *Cancer Institute NSW Early Career Fellowship*, *Cancer Australia Priority-driven Collaborative Cancer Research Grant APP1060919* and a *NHMRC Australia Fellowship*.