Evaluating the Accuracy of 4D-CT Ventilation Imaging: First Comparison with Technegas SPECT Ventilation

Short title: A comparison of 4D-CT Ventilation Imaging and Technegas V-SPECT

Fiona Hegi-Johnson[1,2,3,4]; Paul Keall[1]; Jeff Barber[5]; Chuong Bui[6]; John Kipritidis[1,7]

1. Radiation Physics Laboratory, Faculty of Medicine, Sydney University, Camperdown, NSW, Australia 2006
2. Department of Medical Physics, School of Mathematical and Physical Sciences, University of Newcastle, Newcastle, NSW, Australia, 2300
3. Radiation Oncology Centre, Seventh Day Adventist Hospital, Wahroonga, NSW, Australia, NSW 2076
4. Department of Radiation Oncology, Sir Peter MacCallum Department of Oncology, University of Melbourne, Victoria 3000.
5. Crown Princess Mary Cancer Care Centre, Blacktown Hospital, Blacktown, NSW, Australia, 2148
6. Department of Nuclear Medicine, Nepean Hospital, Nepean, NSW, Australia 2750
7. Department of Radiotherapy, Royal North Shore Hospital, St Leonards 2065.

Corresponding Author:

Dr Fiona Hegi-Johnson

Radiation Physics Laboratory

Blackburn Building

University of Sydney

NSW 2006

Phone: +61 2 9351 2222

Email: fionahegi@gmail.com or Fiona.Hegi-Johnson@petermac.org

Abstract

Introduction:

Computed tomography ventilation imaging (CTVI) is a highly accessible functional lung imaging modality that can unlock the potential for functional avoidance in lung cancer radiation therapy. Previous attempts to validate CTVI against clinical ventilation single-photon emission computed tomography (V-SPECT) have been hindered by radioaerosol clumping artifacts. This work builds on those studies by performing the first comparison of CTVI with ^{99m}Tc-carbon ('Technegas'), a clinical V-SPECT modality featuring smaller radioaerosol particles with less clumping.

Methods:

11 lung cancer radiotherapy patients with early stage (T1/T2N0) disease received treatment planning four-dimensional CT (4DCT) scans paired with Technegas V/Q SPECT/CT. For each patient, we applied three different CTVI methods. Two of these used deformable image registration (DIR) to quantify breathing induced lung density changes ($CTVI_{DIR\text{-}HU}$), or breathing induced lung volume changes ($CTVI_{DIR\text{-}Jac}$) between the 4DCT exhale/inhale phases. A third method calculated the regional product of air-tissue densities ($CTVI_{HU}$) and did not involve DIR. Corresponding CTVI and V-SPECT scans were compared using the Dice Similarity Coefficient (DSC) for functional defect and non-defect regions, as well as the Spearman correlation *r* computed over the whole-lung. The DIR target registration error (TRE) was quantified using both manual and computer selected anatomic landmarks.

Results:

Interestingly the overall best performing method ($CTVI_{HU}$) did not involve DIR. For non-defect regions, the $CTVI_{HU}$, $CTVI_{DIR\text{-}HU}$, and $CTVI_{DIR\text{-}Jac}$ methods achieved mean DSC values of 0.69, 0.68, and 0.54 respectively. For defect regions, the respective DSC values were moderate: 0.39, 0.33 and 0.44. The Spearman *r* values were generally weak: 0.26 for

$CTVI_{HU}$, 0.18 for $CTVI_{DIR-HU}$, -0.02 and for $CTVI_{DIR-Jac}$. The spatial accuracy of CTVI was not significantly correlated with TRE, however the DIR accuracy itself was poor with TRE > 3.6 mm on average, potentially indicative of poor quality 4DCT. Q-SPECT scans achieved good correlations with V-SPECT (mean r >0.6), suggesting that the image quality of Technegas V-SPECT was not a limiting factor in this study.

Conclusion:

We performed a validation of CTVI using clinically available 4DCT and Technegas V/Q-SPECT for 11 lung cancer patients. The results reinforce earlier findings that the spatial accuracy of CTVI exhibits significant inter-patient and inter-method variability. We propose that the most likely factor affecting CTVI accuracy was poor image quality of clinical 4DCT.

1. Introduction

CT ventilation imaging (CTVI) combines respiratory correlated four-dimensional CT (4D-CT) with deformable image registration (DIR) to visualize breathing-induced air volume changes in the lung [1,2]. As 4D-CT is increasingly considered standard of care for treatment planning in lung cancer radiotherapy, CTVI provides "free information" permitting an individualised approach to the planning of lung cancer radiotherapy[3,4,5]. In 2016 CTVI-guided functional avoidance was applied clinically for the first time[6], however further work is still needed at the basic level to quantify the spatial accuracy of CTVI.

The clinical gold standard for assessing regional lung function is ventilation / perfusion single photon emission computed tomography (V/Q-SPECT) using inhaled and injected radioisotopes, namely ^{99m}Tc-labeled diethylenetriamine pentacetate (DTPA) and macroaggregated albumin (MAA). Previous attempts to validate CTVI using DTPA V-SPECT have indicated weak spatial accuracy (voxel-level correlations in the range 0.1-0.4), and this is partly attributed to focal clumping of DTPA in the main airways[7,8,9]. By comparison, validation of CTVI against positron emission tomography using ^{68}Ga –labelled nanoparticles ('Galligas PET') has led to improved voxel-level correlations (in the range 0.4-0.5) owing to the smaller particle size of Galligas compared to DTPA[10,11,12]. The main drawback of Galligas is that it is considered an experimental modality, and the specialised requirements for Galligas generation limit the opportunities for larger scale clinical validation of CTVI.

The purpose of this study was to perform the first evaluation of CTVI using a different V-SPECT modality based on ^{99m}Tc- Carbon ('Technegas')[13,14,15]**.** Like Galligas, Technegas is a smaller molecule than DTPA and disperses throughout normal lung with less clumping and no washout. For the purposes of widespread validation of CTVI, Technegas has the additional advantage of being commercially available internationally. In this work, we

replicate the analyses of previous CTVI validation studies using Technegas V-SPECT. Specifically, we evaluate the Dice similarity coefficient (DSC) for both ventilation defect and non-defect regions and the Spearman correlation *r* evaluated across the whole lung. We test the three main classes of CTVI present in the literature: two of which use DIR to evaluate breathing-induced changes in lung volume or density as visible in 4DCT. A third method uses the CT number to estimate the regional product of air-tissue densities without DIR.

A major challenge for CTVI validation is that there can exist large variations in CTVI accuracy between different subjects and different CTVI methods (see for example Figure 1, which exhibits the best and worst patient cases from this study). To better characterize this, we perform a number of analyses beyond those performed in previous studies. Namely, in addition to correlating CTVI with V-SPECT, we also correlate CTVI against the corresponding Q-SPECT scans to determine if the V-SPECT image quality was a limiting factor. We additionally investigate the influence of DIR accuracy as quantified by the target registration error (TRE) for both manual- and computer-selected anatomic landmarks, and consider the impact of time-delays between the 4DCT and V/Q-SPECT scans by generating CTVIs directly from the V/Q-SPECT localization CT. Finally, we calculate the correlations between V-SPECT and Q-SPECT scan directly, which is anticipated to represent an “upper bound” on the CTVI accuracy.

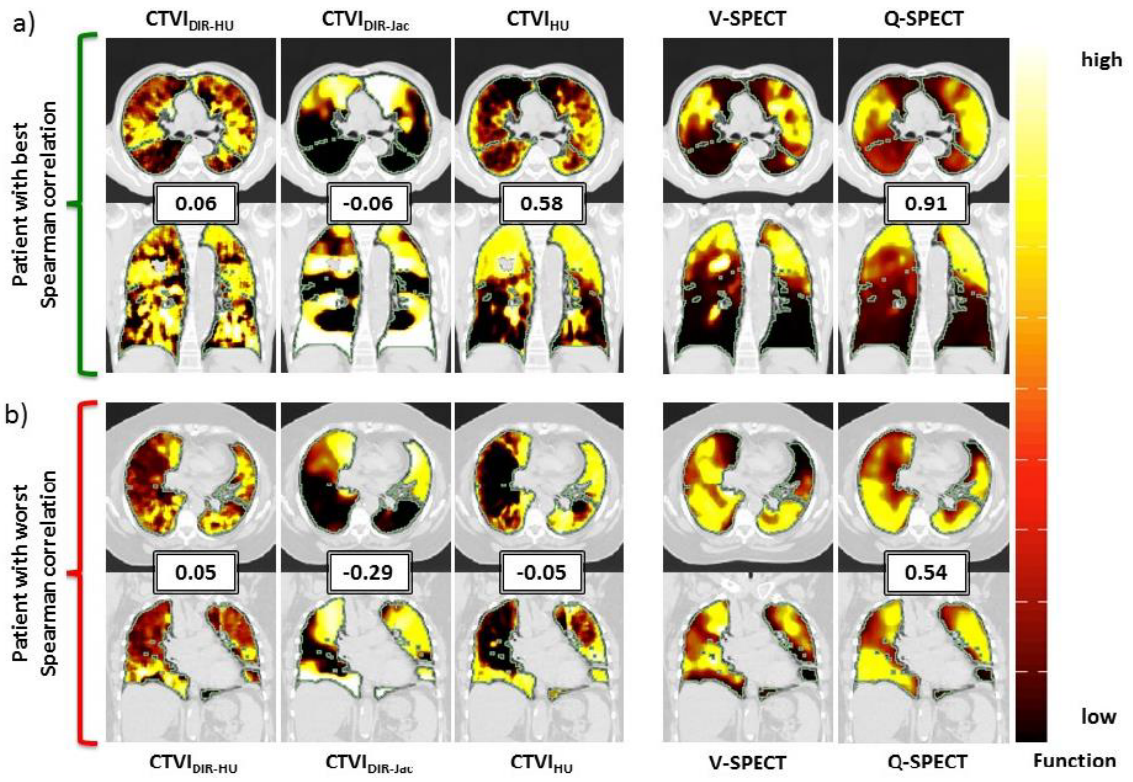


Figure 1: Illustrating inter-patient and inter-method variability in CTVI as compared to V/Q-SPECT. In each panel the functional image is overlaid on the time-averaged 4DCT. These two patients were selected to include: (a) the single best case, and (b) the single worst case of Spearman correlation between CTVI and V-SPECT in this study. In each panel the number represents the Spearman correlation with V-SPECT. See text for details on the 4DCT and V/Q-SPECT acquisitions, and computation of the different CTVI methods (denoted by subscripts "DIR-HU," "DIR-Jac" and "HU").

2. Methods and Materials

2.1. Study Design

11 patients were drawn from a prospective, single-arm, ethics-approved clinical trial through the Western Sydney Local Health District (clinical trial number ACTRN12614000478617). Patients were eligible if they had early stage primary non-small

cell lung cancer and were suitable for treatment with stereotactic ablative body radiotherapy (SABR). All patients were ≥ 18 years of age and provided written informed consent.

Table 1: Patient characteristics

Characteristic	Value (%)
Age (years) (mean +/- SD)	77±8
Sex Male Female	 4 (36) 7 (64)
Time between 4D-CT and V/Q SPECT (days) mean (range)	33 (1-95)
Tumour location by lobe RUL/RML RLL LUL LUL	 5 (45) 2 (18) 4 (36) 0 (0)
Central vs. Peripheral Zone Central Zone Peripheral Zone	 3(27) 8(73)
Dose 48Gy/4 50 Gy/5	 9 (81) 2 (19)
PFT's (mean +/- SD) FEV_1* (% pred) FEV_1/FVC * (% pred) DL_{C0}+ (% pred)	 61±26 71 ± 25 52 ± 11
Abbreviations: DL_{CO} = diffusing capacity of the lung for carbon monoxide; FEV_1 = forced expiratory volume in 1 second; FVC = forced vital capacity * available for 6 patients + available for 5 patients	

125 Patients underwent radiotherapy treatment planning with 4D-CT and were assessed with V/Q SPECT. Seven patients had V/Q SPECT images acquired before treatment and 4 after radiotherapy treatment had commenced. All patients had inoperable lung cancer, and the

majority had significant impairment of respiratory function based on pulmonary function tests (PFTs) performed before treatment. The patient characteristics are shown in Table 1.

2.2. Details of the 4D-CT, Technegas V/Q-SPECT and PFT examinations

Each patient underwent a treatment planning 4DCT scan with a GE Lightspeed RT 16- slice scanner (GE Medical Systems, Waukesha, WI). Respiratory monitoring was performed using the Varian RPM system (Varian Medical Systems, Palo Alto, CA) with the 4DCT reconstructed into 10 phase bins using the Advantage 4D software (GE Healthcare). The 4DCT scans had 512×512 pixels with pixel size 1×1 mm^2 and slice thickness 2 mm.

V/Q SPECT projections and low-dose CT scans for attenuation correction were acquired using a Philips Brightview XCT camera. Technegas was administered prior to acquiring the V-SPECT, with patients instructed to take slow, deep breaths to maximise dispersal of the aerosol in the pulmonary parenchyma. Technetium macro-aggregated albumin (Tc99m-MAA) was then administered intravenously followed by Q-SPECT acquisition.

V/Q SPECT scans comprised 64 projections with acquisition times of 10 seconds per projection for V-SPECT and 8 seconds per projection for Q-SPECT during tidal breathing. Projections were reconstructed into a 128 × 128 matrix of pixel size 4.7 mm and slice spacing 4.7 mm using Astonish iterative reconstruction (4 iterations and 8 subsets).

2.3. Alignment and segmentation of the 4DCT and V/Q-SPECT scans

The V/Q-SPECT images were rigidly aligned to the 4DCT exhale phase image using the low-dose SPECT/CT in Velocity AI (Varian Medical Systems). As a result of this alignment procedure the V/Q-SPECT scans were linearly resampled to the 4DCT voxel spacing of 1×1×2 mm^3. As in Ref. [11] a median filter of kernel width 7×7×7 $voxels^3$

($9\times9\times18$ mm^3) was applied to all V/Q-SPECT images to minimize the influence of image noise.

The delineation of lung lobes in 4DCT is challenging as a result of irregular-breathing induced truncation/duplication artefacts and also because the fissure width is thinner than the 4DCT slice thickness. To minimize this problem, we delineated the region of interest (ROI) for each CTVI and V/Q-SPECT comparison on the 4DCT exhale phase image, as this is the most stable in terms of image quality. However, in several cases the right middle lobe boundary was still difficult to see. Therefore, the lung was divided into the following regions: left upper lobe (LUL), left lower lobe (LLL), right lower lobe (RLL) and the right upper lobar region (RULR), which included both the right upper and right middle lobes.

For each patient we then defined the whole lung ROI by taking the union of the LUL, LLL, RLL and RULR regions.

2.4. CTVI generation

CTVIs were generated from the 4DCT scans using VESPIR (*VEntilation via Scripted Pulmonary Image Registration*)[16] which was previously used to compare CTVI against Galligas PET[11] and we apply the same CTVI algorithm parameters here. Briefly, the method performs a B-spline DIR between each adjacent pair of 4DCT phase images (e.g. we deform Phase 2→ 1, Phase 3→ 2, Phase 4→ 3, and so on), and respectively each individual DIR operation produces a motion field pointing from Phase 1→2, Phase 2→3, Phase 3→ 4, etc. As in Kipritidis et al.[17] we then filtered out the error from each individual DIR process by assuming that the composed motion field over the whole breathing cycle should add to zero. Finally, we composed the corrected motion field between the exhale and inhale phase images which is taken as the motion associated with ventilation. The DIR used an intensity mean square error (MSE) similarity metric with a scalar regularization parameter λ=1 to ensure

spatial smoothness of the DIR motion fields. We performed an initial visual check of the DIR results by comparing the alignment of lung structures between the 4DCT exhale/inhale phase images both before and after DIR.

Three types of CTVIs were then created based on different ventilation surrogates: (i) breathing induced lung density change ($CTVI_{DIR\text{-}HU}$), (ii) breathing induced lung volume change ($CTVI_{DIR\text{-}Jac}$) and (iii) the regional product of air and tissue densities ($CTVI_{HU}$). The CTVIs were calculated by evaluating the following expressions at each voxel location *x*,

1) $\mathrm{CTVI}_{\mathrm{DIR\text{-}HU}}(x) = \frac{[\mathrm{HU}_{\mathrm{ex}}(x) - \mathrm{HU}^{*}_{\mathrm{in}}(x)]}{\left[\mathrm{HU}^{*}_{\mathrm{in}}(x) + 1000\right]} \times \frac{[\mathrm{HU}_{\mathrm{ex}}(x) + 1000]}{1000}$, where $\mathrm{HU}_{\mathrm{ex}}$ and $\mathrm{HU}_{\mathrm{in}}$ refer to the maximal exhale and registered inhale phases, which is corrected by the mass correction factor $(*)$

2) $\mathrm{CTVI}_{\mathrm{HU}}(x) = \sum_{\varphi=1}^{10}\left[\frac{\mathrm{HU}_{\varphi}(x)}{-1000} \times \frac{[\mathrm{HU}_{\varphi}(x) + 1000]}{1000}\right] / 10$ where HUφ(x) is the HU value at voxel location (x) and 4DCT phase bin φ = 1,. . .,N.

3) $\mathrm{CTVI}_{\mathrm{DIR\text{-}Jac}}(x) = [\mathrm{Jac}(x) - 1]$ where Jac(x) is the Jacobian determinant of deformation.

Respectively, the $CTVI_{DIR\text{-}HU}$ and $CTVI_{DIR\text{-}Jac}$ methods rely on DIR to evaluate regional Hounsfield Unit (HU) changes [1] or to calculate the Jacobian determinant of deformation ('Jac') describing regional volume change [2]. These represent the two dominant forms of CTVI in the literature. The third method ($CTVI_{HU}$) is a streamlined approach that incorporates HU information from across the whole 4D cycle and does not rely on DIR[11]. $CTVI_{HU}$ methods are more sensitive to motion blurring than DIR based methods, with the breathing motion directly related to the spatial extent of the blurring; we would generally expect higher blurring at the diaphragm compared to the apex. It is also worth noting that the SPECT V/Q images themselves suffer from motion blur as they were acquired under free-breathing without gating. All CTVIs were normalized by the 90th percentile of ventilation

inside the lung, and a median filter of kernel width 7×7×7 voxels^3 (9×9×18 mm^3) was applied to minimize the influence of small scale DIR errors and image noise.

2.5. Segmentation of ventilation / perfusion defect and non-defect regions

For V/Q-SPECT images, defect regions were segmented using an image-specific intensity threshold set at 50% of the 90th percentile within the whole lung ROI. This algorithmic approach was used in an earlier Galligas-PET study[11] and provided good agreement between clinician and computer selected thresholds. In this study the computer-segmented V/Q-SPECT defect regions were visually reviewed by one of the authors.

For CTVI, there exists is no consensus on the best thresholding method. Rather we tested a number of different possible defect intensity thresholds set at 5% increments (i.e. 5%, 10% and so on up to 95%) of the 90th percentile of ventilation within the whole lung ROI. For each different CTVI type, we optimised the threshold to the nearest 5% across the whole patient population by minimising the residual of non-defect lung volumes between CTVI and V-SPECT. The resulting thresholds for the $CTVI_{DIR\text{-}HU}$, $CTVI_{DIR\text{-}Jac}$ and $CTVI_{HU}$ methods were selected as 20%, 30% and 70% of the 90th percentile ventilation, respectively.

2.6. Voxel-based comparisons of CTVI and V/Q-SPECT

2.6.1. Dice similarity coefficient for functional defect and non-defect regions

The Dice similarity coefficient (DSC) describes the fractional volume overlap between two regions (in our case, ventilation/perfusion defect regions or non-defect regions) and takes a value in the range [0,1]. For example, the DSC for defect regions in CTVI and V-SPECT was calculated using,

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

where the notation "|A|" denotes the volume of a region A, and "|A ∩ B|" indicates the volume of the intersection of regions A and B. We similarly calculated DSC values for non-defect regions.

2.6.2. Spearman correlations in whole lung region of interest

The whole-lung Spearman correlation *r* was computed between each different CTVI type and corresponding Technegas V/Q-SPECT scans. The Spearman *r* values are defined in the range [−1, 1] and indicate the degree of monotonicity of values in spatially matched voxels within the whole lung ROI.

2.6.3. Assessing the Impact of DIR performance

We assessed the dependence of CTVI accuracy on DIR performance by calculating the target registration error (TRE) for a set of anatomic landmark pairs defined on each 4DCT exhale and inhale phase image pair. The DIR motion field was then used to warp the inhale-landmarks to the exhale geometry in order to calculate TRE. Two independent landmark selection methods were applied: one was a semi-automated (or 'manual') approach and the other was a fully-automated method. For the manual approach, one of the authors selected up to 50 intensity based landmark pairs in the lung parenchyma using the Utrecht iX landmark tool[18]. A second author then reviewed and corrected each of these landmarks where appropriate. The fully-automated method used the scale invariant feature transform (SIFT) method as implemented by Paganelli et al.[19]. The SIFT algorithm produced in excess of 100 landmark pairs in the lung parenchyma for each 4DCT scan. As in our previous studies[10] the final TRE for both landmark selection methods, excluded any landmark pairs where the (pre-

DIR) landmark distance was more than 2.5 standard deviations outside the mean landmark displacement for that patient.

2.6.4. Additional cross-modality comparisons using Q-SPECT

We considered it reasonable to assume that the V/Q SPECT scans themselves should be well correlated, as all of the V/Q SPECT scans in our study were reviewed by a nuclear medicine physician and no notable V/Q mismatches were found. Given also that Q-SPECT suffers less noise than V-SPECT, and is acquired in short succession after V-SPECT, we computed the DSC and Spearman *r*-values between corresponding CTVI and Q-SPECT scan pairs, and also between V-SPECT and Q-SPECT scan pairs to determine if this could produce improved cross-modality correlations.

2.6.5. Assessing the Impact of time-delays between 4DCT and V/Q-SPECT

In order to further assess the possible influence of time delays between corresponding 4D-CT and V/Q SPECT scans (which ranged between 1 and 95 days; see Table I), we calculated the linear (Pearson) correlation between the time delay in days and the Spearman *r* values between CTVI and V-SPECT. To overcome the possible influence of time-delays on the accuracy of CT ventilation, we additionally computed the DSC and Spearman *r*-values between each V-SPECT and its corresponding low-dose CT scan with the $CTVI_{HU}$ method applied.

2.7. Lobar-level comparisons of CTVI and V/Q SPECT

In addition to the voxel level comparisons of Sec. 2.6. it is also of interest to consider the accuracy of CTVI at a coarser level of spatial resolution. Similar to the study by Eslick et al.[12] we compared CTVI and V/Q SPECT in terms of the contribution of each lobe to the total ventilation for that patient. This was achieved by computing the sum of CTVI values in each manually delineated lobar region (LLL, LUL, RLL and RULR) and dividing by the sum

of ventilation values in the whole lung ROI. We calculated the linear (Pearson) correlations between CTVI- and SPECT-derived lobar contributions for 44 lobar regions across the 11 patients.

2.8. Global comparisons (Coefficient of Variation)

For each CTVI or SPECT image, the Coefficient of Variation (CoV) was computed by taking the ratio of the standard deviation to the mean of all voxel values within the whole lung ROI. The CoV provides a measure of overall image heterogeneity for the functional distribution and is expected to vary proportionally between CTVI and V/Q-SPECT. For each corresponding combination of CTVI and SPECT image, we calculated the linear (Pearson) correlation of CoV values across all 11 patients.

3. Results

3.1. Voxel-based comparisons of CTVI and V/Q-SPECT

3.1.1. Dice similarity coefficient for functional defect and non-defect regions

Figure 2(a) shows the DSC values obtained for non-defect regions between each of the $CTVI_{DIR\text{-}HU}$, $CTVI_{DIR\text{-}Jac}$ or $CTVI_{HU}$ methods and the corresponding V-SPECT or Q-SPECT scans across all 11 patients. Comparing CTVI with V-SPECT (white boxes), the $CTVI_{DIR\text{-}HU,}$ $CTVI_{DIR\text{-}Jac}$ and $CTVI_{HU}$ methods achieved (mean ± SD) DSC values of (0.68 ± 0.54), (0.54 ± 0.13) and (0.69 ± 0.08), respectively. When comparing CTVI against Q-SPECT (light shaded boxes), the respective DSC values were slightly higher: (0.74 ± 0.14), (0.60 ± 0.14) and (0.76 ± 0.07). The best cross-modality agreement was between V-SPECT and Q-SPECT (dark shaded boxes) which had (mean ± SD) values of (0.81 ± 0.05).

Similarly Figure 2(b) shows the DSC values for defect regions; here the accuracy of CTVI was only moderate. Comparing against V-SPECT, the $CTVI_{DIR\text{-}HU,}$ $CTVI_{DIR\text{-}Jac}$ and $CTVI_{HU}$ methods achieved DSC values of (0.33 ± 0.15), (0.44 ± 0.17) and (0.39 ± 0.18) respectively.

The agreement between defect regions in CTVI and Q-SPECT was similar, with respective DSC values of (0.35 ± 0.15), (0.43 ± 0.14) and (0.41 ± 0.20). Once again the best observed agreement was between V-SPECT and Q-SPECT, with mean DSC: (0.67 ± 0.15).

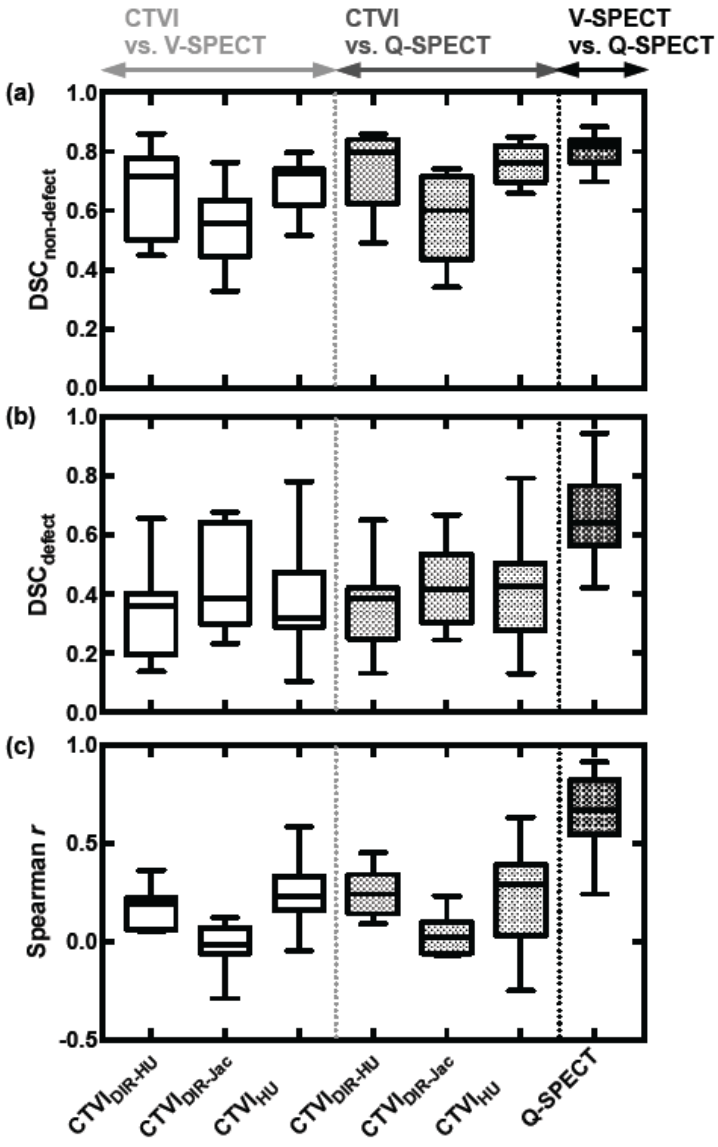


Figure 2: Boxplots showing correlation values between different CTVI methods and the corresponding V/Q-SPECT scans. Higher correlations indicate better CTVI accuracy. The panels show (a) DSC evaluated for non-defect regions, (b) DSC evaluated for defect regions,

and (c) Spearman *r*-values evaluated over the whole lung. In each panel, the white boxes refer to comparisons between CTVI and V-SPECT, light shaded boxes refer to comparisons between CTVI and Q-SPECT and the dark shaded boxes refer to comparisons between V-SPECT and Q-SPECT. For each box, the upper, middle and lower edges of the box represent the upper quartile, median and lower quartile of r values over all 11 patients.

3.1.2. Spearman correlations evaluated across the whole lung

Figure 2(c) compares the voxel-wise Spearman correlations obtained between each of the $CTVI_{DIR\text{-}HU}$, $CTVI_{DIR\text{-}Jac}$ and $CTVI_{HU}$ methods and V-SPECT or Q-SPECT for all patients. When comparing CTVI and V-SPECT, the Spearman *r* values were generally weak, with (mean ± SD) values of (0.18 ± 0.10) for $CTVI_{DIR\text{-}HU}$, (-0.02 ± 0.11) for $CTVI_{DIR\text{-}Jac}$ and (0.26 ± 0.18) for $CTVI_{HU.}$ The performance of the DIR-based CTVIs was slightly improved when comparing against Q-SPECT; (0.24 ± 0.12) for $CTVI_{DIR\text{-}HU}$ and (0.03 ± 0.09) for $CTVI_{DIR\text{-}Jac}$. The correlations between $CTVI_{HU}$ with Q-SPECT were not much different to the comparison with V-SPECT (0.24 ± 0.25). By far the best observed Spearman correlations were those calculated between V-SPECT and Q-SPECT directly, with (mean ± SD) values of (0.66 ± 0.19). For all but one patient, the correlation between V-SPECT and Q-SPECT exceeded the correlation between V- or Q-SPECT with any of the CTVIs.

3.1.3. Assessing the Impact of Time Delays between 4DCT and V/Q SPECT scans

We observed no significant link between the Spearman correlation values and the time delay between the 4DCT and SPECT scans; the Pearson correlation values were -0.13 (p=0.69), -0.29 (p=0.39) and 0.13 (p=0.71) for the $CTVI_{\_DIR\text{-}HU}$, $CTVI_{\_DIR\text{-}Jac}$ and $CTVI_{\_HU}$ methods respectively. In fact, both the highest and lowest Spearman correlation between any CTVI and its corresponding V-SPECT scan occurred for time delays of 69 and 60 days

respectively; see Figure 1(a) and (b) respectively. Applying the $CTVI_{HU}$ method to the low-dose CT, acquired on the same day as the SPECT scans, and comparing this with the corresponding V-SPECT scans, we found that the (mean +\- SD) Spearman correlation was (0.12 +\- 0.18), poorer than for the case of the $CTVI_{HU}$ method as applied to the 4DCT scans. This suggestive that the time delay between 4DCT and V-SPECT scans was not the dominant source of error in our analysis.

3.1.4. Assessing the Impact of DIR performance

Based on our analysis of the TRE both before and after DIR, we conclude that the DIR performance in this study was relatively poor. Our manual landmarking method produced between 42-48 landmarks for each patient, resulting in a (mean ± SD) TRE of (6.58 ± 2.58) mm before DIR and (4.44 ± 1.18) mm after DIR. The fully-automated SIFT method produced between 150-417 landmarks for each patient and resulted in similar TRE of (6.26 ± 2.24) mm before DIR and (3.62 ± 1.33) mm after DIR.

In the case of the manually selected landmarks, the % reduction in TRE after DIR was positively correlated with the accuracy of $CTVI_{DIR\text{-}HU}$ method, as compared with V-SPECT and measured using the Spearman r-values. In this particular case the linear correlation was 0.52 ($p=0.10$). Aside from this, no other statistically significant correlations were observed between TRE and the Spearman-based assessment of CTVI accuracy.

3.2. Lobar-based comparisons of CTVI and V/Q SPECT

Figures 3(a) –(d) show the correlations between $CTVI_{DIR\text{-}HU}$, $CTVI_{DIR\text{-}Jac}$ $CTVI_{HU}$ and Q-SPECT versus V-SPECT in terms of the contribution of each lobe to the total ventilation for that patient. Here $CTVI_{HU}$ performed the best out of all the CTVI methods, exhibiting good agreement with Technegas V-SPECT (Pearson correlation 0.79, $p<0.001$). Moderate to strong agreement was also obtained for the DIR based methods; the Pearson correlation for

$CTVI_{DIR\text{-}HU}$ was 0.54 ($p<0.001$) and for $CTVI_{DIR\text{-}Jac}$ it was 0.45 ($p=0.002$). The comparison between V-SPECT vs. Q-SPECT showed strong agreement at the lobar level with a linear correlation of 0.89 ($p<0.0001$).

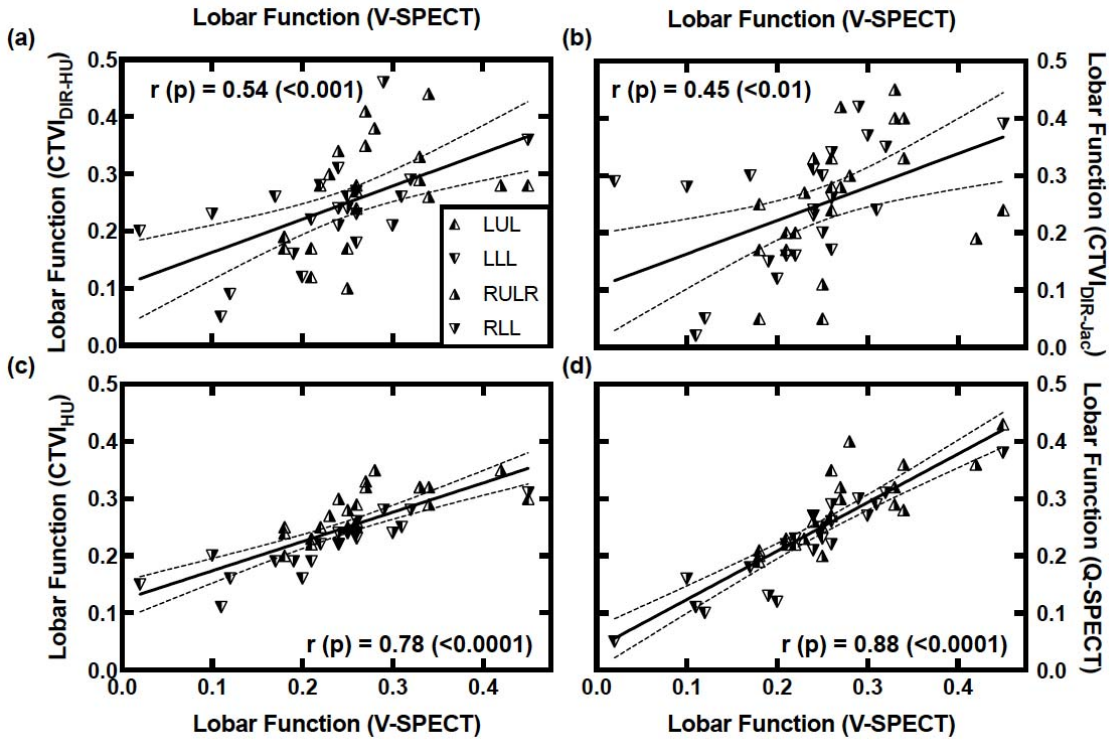


Figure 3: Comparing different functional lung images with V-SPECT in terms of the contribution of different lobar regions to the total function for that patient. The subpanels compare V-SPECT with: (a) $CTVI_{DIR\text{-}HU}$, (b) $CTVI_{DIR\text{-}Jac}$, (c) $CTVI_{HU}$ and (d) Q-SPECT. As a guide to interpreting this figure, upper and lower lobes are represented with triangles facing up or down, respectively. Similarly left and right-sided lobes are represented with triangles that are, respectively, filled on the left and right sides. In each panel, the solid lines (dashed curves) show the linear regression (95% confidence interval). The given r (p) values refer to the linear correlation.

3.3. Global comparisons (Coefficient of Variation)

The image heterogeneity was assessed by calculating the CoV, which had (mean±SD) values of 1.16 ± 0.96, 1.46 ± 0.71 and 0.20 ± 0.06 for $CTVI_{DIR\text{-}HU,}$ $CTVI_{DIR\text{-}Jac}$ and $CTVI_{HU}$ respectively. The V and Q-SPECT images had CoV being 0.66±0.34 and 0.49±0.19 for V-SPECT and Q-SPECT respectively.

Figure 4 compares CoV between V-SPECT and each of the different CTVI methods (as well as Q-SPECT). Based on CoV values generated from all 11 patients, the linear correlation values were 0.24 (p=0.48) for $CTVI_{DIR\text{-}HU}$, 0.43 (p=0.18) for $CTVI_{DIR\text{-}Jac}$ and 0.78 (p<0.01) for $CTVI_{HU}$. As was the case in Sec. 3.1. and 3.2., $CTVI_{HU}$ was the best performing CTVI method, however in this case $CTVI_{DIR\text{-}Jac}$ appears more accurate than $CTVI_{DIR\text{-}HU}$. The best agreement was between the V- and Q-SPECT scans, with a linear correlation 0.88 (p<0.001).

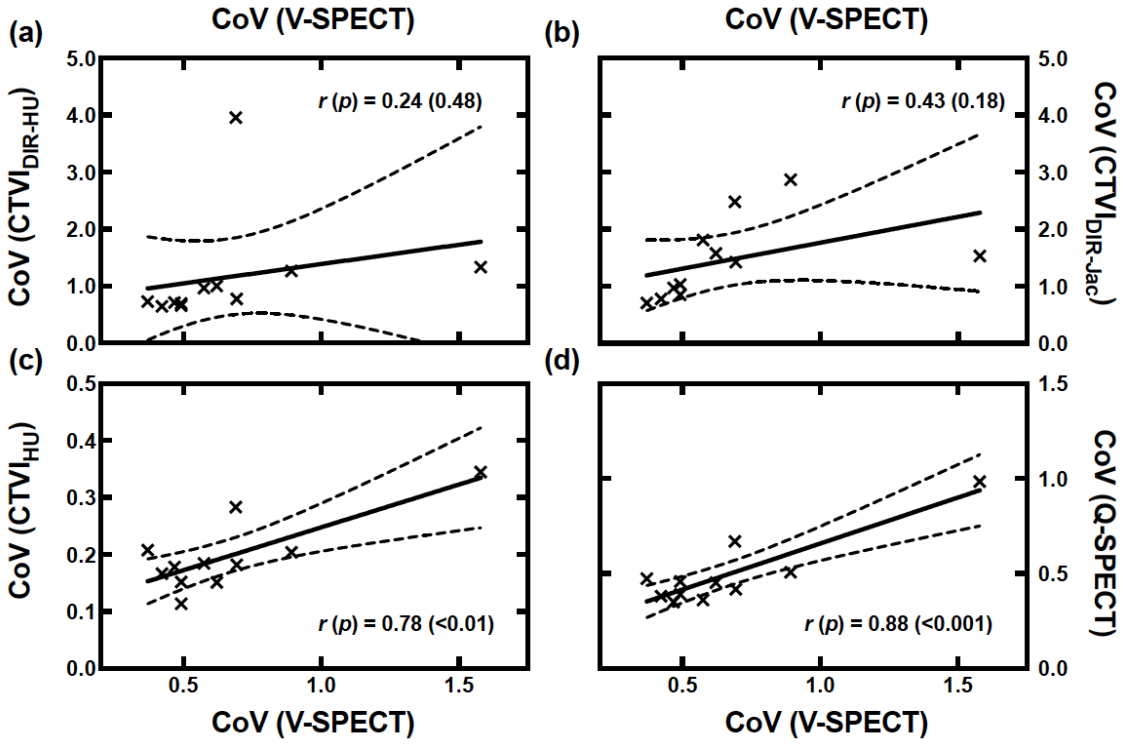


Figure 4: Comparing different functional lung image types with V-SPECT in terms of the coefficient of variation (CoV). The subpanels compare V-SPECT with: (a) $CTVI_{DIR-HU}$, (b) $CTVI_{DIR-Jac}$, (c) $CTVI_{HU}$ and (d) Q-SPECT. In each panel, the solid lines (dashed curves) show the linear regression (95% confidence interval). The given r (p) values refer to the linear correlation.

4. Discussion

This study describes the first validation of CTVI using clinically available Technegas V-SPECT, which is important as it can inform more widespread validation of CTVI in the future. This study reinforces the findings of earlier studies that the spatial accuracy of CTVI can vary from patient-to-patient, as well as with the choice of CTVI method and the metric used to evaluate CTVI accuracy. Compared to V-SPECT, the overall best performing CTVI method ($CTVI_{HU}$) achieved good DSC values for non-defect regions (mean value 0.68) and moderate DSC values for defect regions (mean value 0.39). However, the Spearman correlations between CTVI and V-SPECT evaluated across all lung voxels were relatively weak (the $CTVI_{HU}$ method had a mean *r*-value of 0.26). The accuracy of CTVI appears stronger when evaluated at more coarse levels of spatial resolution: for example the $CTVI_{HU}$ method showed very good agreement with V-SPECT in terms of the per-lobe contribution to total lung function (linear correlation 0.79) and in terms of the CoV (linear correlation 0.78).

Table 2: Comparison of this study with other V-SPECT Validation Papers

Study	V-SPECT modality	Patient number	CTVI metric	DSC (defect)	DSC (non-defect)	Spearman *r*
Yamamoto (2010)[8]	DTPA	1	$CTVI_{DIR-HU}$ $CTVI_{DIR-Jac}$	- -	- -	0.03 0.18
Castillo (2010)[7]	DTPA	7	$CTVI_{DIR-HU}$ $CTVI_{DIR-Jac}$	0.35 0.32	≥ 0.2 ≥ 0.2	- -

Kida (2016)[5]	DTPA	8	$CTVI_{DIR-HU}$ $CTVI_{DIR-Jac}$	- -	- -	0.47 0.37
This study	Technegas	11	$CTVI_{DIR-HU}$ $CTVI_{DIR-Jac}$ $CTVI_{HU}$	0.33 0.44 0.39	0.68 0.54 0.69	0.18 -0.02 0.26

It is instructive to put our results into perspective; **Table 2** compares our DSC values and Spearman correlations with previous studies [5, 7, 8] comparing CTVI against V-SPECT at the voxel level. We note it is difficult to compare these studies directly, owing to differences in the patient cohorts, patient breathing manoeuvres, 4DCT imaging protocols, CTVI algorithms, ventilation image post-processing and in the time delay between 4DCT and V-SPECT scans. Most of these earlier studies demonstrate correlations in the range (0-0.5) and our observations are largely consistent with that. To date, the highest Spearman correlations between CTVI and V-SPECT remain those reported by Kida et al.[5] who obtained Spearman r=0.44 averaged over 8 patients who were imaged with DTPA, but with non-severe clumping. The Kida study showed that functional avoidance treatment plans derived from CTVI and DTPA V-SPECT can exhibit comparable functional dosimetry despite only moderate CTVI accuracy. Even so, more accurate CTVI would be desirable, particularly for the case highly targeted radiotherapy treatments such as SABR.

An important component of our work was to perform a series of sub-studies to better characterize the variability of CTVI accuracy. One limitation of our study was the large time-delay between 4DCT and V/Q-SPECT scans, which had a mean value of 33 days (range 1-95 days). Notably however, we did not observe any statistically significant correlations between the length of this time delay and the agreement between CTVI and V-SPECT as quantified by the Spearman r-values. We also attempted to overcome the problem of time-delays by applying the $CTVI_{HU}$ method directly to the SPECT localization CT, which represents the "best-case scenario" in terms of anatomic and temporal alignment between CTVI and V-

SPECT. But the $CTVI_{HU}$ derived from the localization CT showed only poor Spearman correlation with V-SPECT (mean *r*-value 0.12, less than was the case for $CTVI_{HU}$ derived from 4DCT). These observations would appear to discount the severity of time-delays between 4DCT and V-SPECT as the leading source of CTVI error in this study.

Similarly, the image quality of Technegas V-SPECT did not appear to be a limiting factor in this study, as the different CTVI methods demonstrated similar correlations with both V-SPECT and Q-SPECT (see Figure 2). This is in contrast to the studies by Castillo et al.[7, 20] where CTVI showed significantly better correlations with Q-SPECT than (DTPA) V-SPECT. In our study the Spearman correlations between corresponding Technegas V-SPECT and Q-SPECT were good (0.66 on average); and this is the level of agreement we would expect for two functional lung imaging modalities that are physiologically correlated.

Our study population differs from some of the previously CTVI validation papers, as we included only SABR patients, who have early stage lung disease. Patients who have early stage disease are likely to have smaller tumors that don't block airways; therefore, it could be that these patients are more likely to have homogenous ventilation images running the risk that we are comparing noise between the two imaging modalities rather than ventilation defects. However, in our patients we still found a high degree of heterogeneity, as seen for the worst case patient in **Figure 1**, who suffered from a large ventilation defect in both lower lobes. Across 11 patients we found CoV values with (mean±SD) values of 1.16 ± 0.96, 1.46±0.71 and 0.20±0.06 $CTVI_{DIR\text{-}HU,}$ $CTVI_{DIR\text{-}Jac}$ and $CTVI_{HU}$ respectively. The $CTVI_{DIR\text{-}HU}$ CoV figures are comparable with the figures of Brennan at al[21], who demonstrated a CoVof 0.83 for poor functioning lung and 0.53 for good functioning lung using a DIR based HU method similar to that used in this study[21].

In addition, we have examined the proportion of lung ventilated and perfused in the V/Q SPECT. Across all eleven patients, the mean ± SD (range) of the percentage of ventilated lung was 53.7±15.5 (26.1-77.6), and perfused lung was 62.8 ±15.2 (45.2-75.2). Ventilated and perfusion scans were thresholded using the same method as CTVI. This corresponds to the work of Vinogradskiy et al. who demonstrated significant ventilation defects in up to 30% early stage lung cancer patients[22]. The wide range supports the fact that there is a great deal of heterogeneity in the images. Furthermore, the % mismatch of ventilation versus perfusion was relatively small with a mean±SD (range) of 7.5±4.7 (2.4-18.5) providing supporting evidence that the V/Q SPECT images should be well correlated for most patients.

By ruling out the other possibilities, we propose that the most likely source of CTVI error in this study was a high prevalence of image artefacts in 4DCT due to irregular breathing. Irregular breathing is known to cause image artefacts – such as anatomic truncation and duplication - in up to 90% of clinical 4DCT scans[3, 23] and 4DCT image quality is also known to impact the reproducibility of DIR-based CTVI methods[23], [24]**.** Poor 4DCT image quality may help to explain the observations of poor DIR accuracy in this study; both manual and automated landmark selection methods suggested a TRE >3.5 mm on average, which could be considered unacceptably large given a slice thickness of 2 mm but is also likely to be representative of most 4D-CT in clinical use.

By comparison the same DIR was found to have a TRE <2 mm in an earlier study using 4D-PET/CT[10]. Interestingly we observed that the CTVI methods relying on DIR ($CTVI_{DIR\text{-}HU}$ and $CTVI_{DIR\text{-}Jac}$) performed less well on average than a method using no DIR at all ($CTVI_{HU}$). Compared to the present study, the earlier study using 4D PET/CT used a lower-dose setting for the 4DCT scan component, a smaller number of phase images (5 vs. 10) and a larger slice thickness (2.5 mm versus 5 mm). Based on the comparison of scan

parameters, we might expect the treatment planning 4DCT scans in the current study to enable more accurate DIR-based CTVI since the scan parameters imply finer spatial / temporal resolution. However, a comparison in terms of scan parameters alone does not account for the potential problems of irregular patient breathing and related motion artifacts in the reconstructed 4DCT phase images.

To understand the influence of 4DCT image artifacts, it is instructive to compare the following subtraction images: between the deformably-registered 4DCT exhale and inhale phase images ($HU_{50\%}$ and $HU^{*}_{0\%}$, left column), between the two phase images around maximal exhale ($HU_{60\%}$ and $HU_{50\%}$, middle column), and between the two phase images around maximal inhale ($HU_{0\%}$ and $HU_{90\%}$, right column). These are presented in Figure 5 below and correspond to the "worst case" patient from Figure 1. Essentially the HU difference between deformably registered exhale/inhale images can be interpreted as the "ventilation signal"; indeed the $CTVI_{DIR\text{-}HU}$ method is directly related to this HU difference distribution via Eq. (1). By comparison, the differences between $HU_{60\%}$,$HU_{50\%}$ and $HU_{0\%}$,$HU_{90\%}$ can be interpreted as "noise", featuring alternating bright/dark bands that are 4 slices thick in the SI direction and corresponding to the abutting couch positions of the cine-mode 4DCT scan. Noting that all panels have the same window/level settings, we observe that HU differences associated with ventilation are barely larger than the HU differences observed between any pair of neighbouring 4D phases. This is a problem because it suggests that substituting the 50% phase with the 60% phase (or similarly the 0% phase with the 90% phase), could lead to severe variations in the resultant CTVI.

It is challenging to quantify the noise in the dynamic HU signal directly. For example, we attempted to quantify the 4DCT image quality in terms of changes in the normalized cross correlation (NCC) between adjacent slice pairs across abutting couch transitions; this is the method suggested by Cui et al.[25] However, in our case we did not observe a significant

correlation between NCC metrics and the Spearman r-values between CTVI and V-SPECT. While the development of new 4DCT image quality metrics is beyond the scope of this study, the problem of poor ventilation signal observed in **Figure 5** was qualitatively observed across all 11 scans in our dataset and is implicated in the poor CTVI accuracy observed in this work.

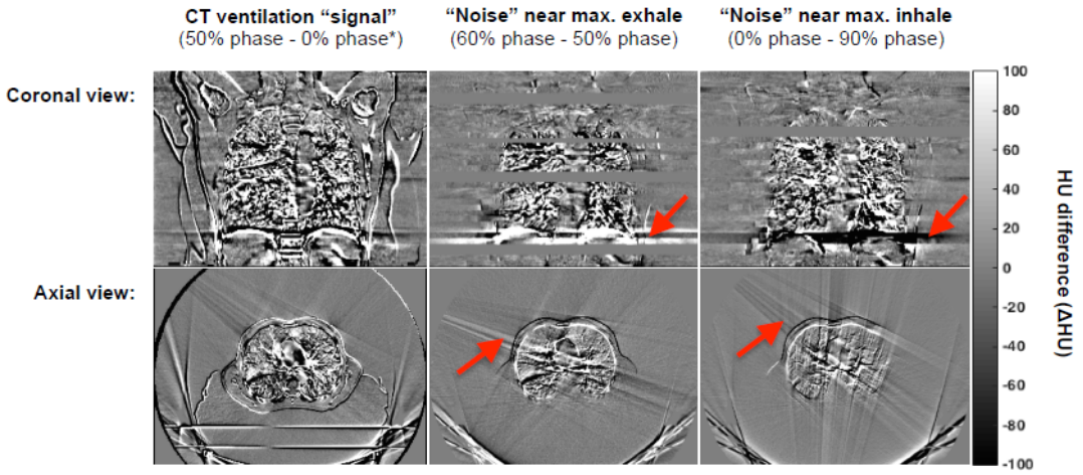


Figure 5: illustrating subtraction images between different 4DCT phase image pairs for the "worst case" patient in Figure 1. Difference images are shown between deformably registered exhale and inhale phase images (left column), the 60%/50% phases around maximum exhale (middle) and the 0%,90% phases around maximum inhale (right). The arrows indicate spurious, artefact induced HU differences outside the lung for 4DCT phase images near exhale (middle column) and near inhale (right columns).

This study represents the validation results that may be expected in a clinical environment, and therefore gives a useful indication of the robustness of CTVI in clinical practice. Technegas appears to be a suitable reference modality, but the quality of 4D-CT itself may have a significant impact on the quality of CTVI, particularly when this is DIR based[26]. In lieu of higher quality 4DCT and/or alternate DIR methods that are robust against stochastic image artifacts, we therefore suggest that the $CTVI_{HU}$ method may prove the most reliable CTVI method for use with clinical 4DCT. Whilst this is still an early result, we aim

to incorporate this dataset into a larger validation dataset in the future, to further investigate the robustness of validation between CTVI and Technegas V-SPECT.

5. Conclusions

Our study compared CTVI with Technegas V-SPECT for 11 lung cancer SABR patients, demonstrating good agreement between CTVI and Technegas V-SPECT in terms of the Dice overlap for non-defect regions, lobar level and whole lung level CoV comparisons. However, the Dice overlap for defect regions, as well as the voxel-wise Spearman correlation showed only weak-moderate agreement. Importantly, the DIR-based CTVI methods performed less well than a method independent of DIR, suggesting a need to optimize the image quality of clinical 4DCT to further improve the accuracy of DIR-based CTVI.

Acknowledgements and Conflicts of Interest

We would like to acknowledge the assistance of the following staff at the Nepean Cancer Care Centre: Dr Peter Flynn (Chair of the Nepean Lung MDT), Dr Roland Yeghiaian-Alvandi, Ms Shamira Cross, Mr Sean White and Ms Katrina West, who assisted with patient recruitment and data acquisition. Dr Kipritidis was supported by a Cancer Institute NSW Early Career Fellowship (13/ECF/1-15). Professor Keall was supported by an NHMRC Australia Fellowship.

References

1. Guerrero T, Sanders K, Castillo E, et al. Dynamic ventilation imaging from four-dimensional computed tomography. *Phys Med Biol*. 2006;51:777-791.

2. Reinhardt JM, Ding K, Cao K, Christensen GE, Hoffman EA, Bodas SV. Registration-based estimates of local lung tissue expansion compared to xenon CT measures of specific ventilation. *Med Image Anal*. 2008;12(6):752-763.

3. Vinogradskiy Y, Castillo R, Castillo E, et al. Use of 4-dimensional computed tomography-based ventilation imaging to correlate lung dose and function with clinical outcomes. *Int J Radiat Oncol Biol Phys*. 2013;86(2):366-371.

4. Yamamoto T, Kabus S, von Berg J, Lorenz C, Keall PJ. Impact of four-dimensional computed tomography pulmonary ventilation imaging-based functional avoidance for lung cancer radiotherapy. *Int J Radiat Oncol Biol Phys*. 2011;79(1):279-288.

5. Kida S, Bal M, Kabus S, et al. CT ventilation functional image-based IMRT treatment plans are comparable to SPECT ventilation functional image-based plans. *Radiother Oncol*. 2016;118(3):521-527.

6. Yamamoto T, Kabus S, Bal M, Keall P, Benedict S, Daly M. The first patient treatment of computed tomography ventilation functional image-guided radiotherapy for lung cancer. *Radiother Oncol*. 2016;118(2).

7. Castillo R, Castillo E, Martinez J, Guerrero T. Ventilation from four-dimensional computed tomography: density versus Jacobian methods. *Phys Med Biol*. 2010;55(16):4661-4685.

8. Yamamoto T, Kabus S, von Berg J, et al. Evaluation of Four-dimensional (4D) Computed Tomography (CT) Pulmonary Ventilation Imaging by Comparison with Single Photon Emission Computed Tomography (SPECT) Scans for a Lung Cancer Patient. In: *Proc. 3rd Int. Workshop on Pulmonary Image Analysis, MICCAI 2010*. ; 2010:117-128.

9. Yamamoto T, Kabus S, Lorenz C, et al. Pulmonary ventilation imaging based on 4-dimensional computed tomography: Comparison with pulmonary function tests and SPECT ventilation images. *Int J Radiat Oncol Biol Phys*. 2014;90(2):414-422.

10. Kipritidis J, Siva S, Hofman MS, Callahan J, Hicks RJ, Keall PJ. Validating and improving CT ventilation imaging by correlating with ventilation 4D-PET/CT using 68Ga-labeled nanoparticles. *Med Phys*. 2014;41(1):011910.

11. Kipritidis J, Hofman MS, Siva S, et al. Estimating lung ventilation directly from 4D CT Hounsfield unit values. *Med Phys*. 2016;43(1):33-43.

12. Eslick EM, Bailey DL, Harris B, et al. Measurement of preoperative lobar lung function with computed tomography ventilation imaging: progress towards rapid stratification of lung cancer lobectomy patients with abnormal lung function. *Eur J Cardiothorac Surg*. 2015:1-8.

13. James JM, Lloyd JJ, Leahy BC, et al. 99TcmTechnegas and krypton-81m ventilation scintigraphy: a comparison in known respiratory disease. *Br J Radiol*. 1992;65(780):1075-1082.

14. Cook G, Clarke SE. An evaluation of Technegas as a ventilation agent compared with krypton-81 m in the scintigraphic diagnosis of pulmonary embolism. *Eur J Nucl Med Mol Imaging*. 1992;19(9):770-774.

15. Jögi J, Jonson B, Ekberg M, Bajc M. Ventilation–perfusion SPECT with 99mTc-DTPA versus Technegas: a head-to-head study in obstructive and nonobstructive disease. *J Nucl Med*. 2010;51(5):735-741.

16. Kipritidis J, Woodruff HC, Eslick EM, Hegi-Johnson F, Keall PJ. New pathways for end-to-end validation of CT ventilation imaging (CTVI) using deformable image registration. In: *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*. IEEE; 2016:939-942. doi:10.1109/ISBI.2016.7493419.

17. Kipritidis J, Hugo G, Weiss E, Williamson J, Keall PJ. Measuring interfraction and intrafraction lung function changes during radiation therapy using four-dimensional cone beam CT ventilation imaging. *Med Phys*. 2015;42(3):1255-1267.

18. Murphy K, van Ginneken B, Klein S, et al. Semi-automatic construction of reference standards for evaluation of image registration. *Med Image Anal*. 2011;15(1):71-84.

19. Paganelli C, Peroni M, Riboldi M, et al. Scale invariant feature transform in adaptive radiation therapy: a tool for deformable image registration assessment and re-planning indication. *Phys Med Biol*. 2012;58(2):287-299.

20. Castillo R, Castillo E, McCurdy M, et al. Spatial correspondence of 4D CT ventilation and SPECT pulmonary perfusion defects in patients with malignant airway stenosis. *Phys Med Biol*. 2012;57(7):1855-1871.

21. Brennan D, Schubert L, Diot Q, et al. Clinical Validation of 4-Dimensional Computed Tomography Ventilation With Pulmonary Function Test Data. *Int J Radiat Oncol Biol Phys*. 2015;92(2):423-429.

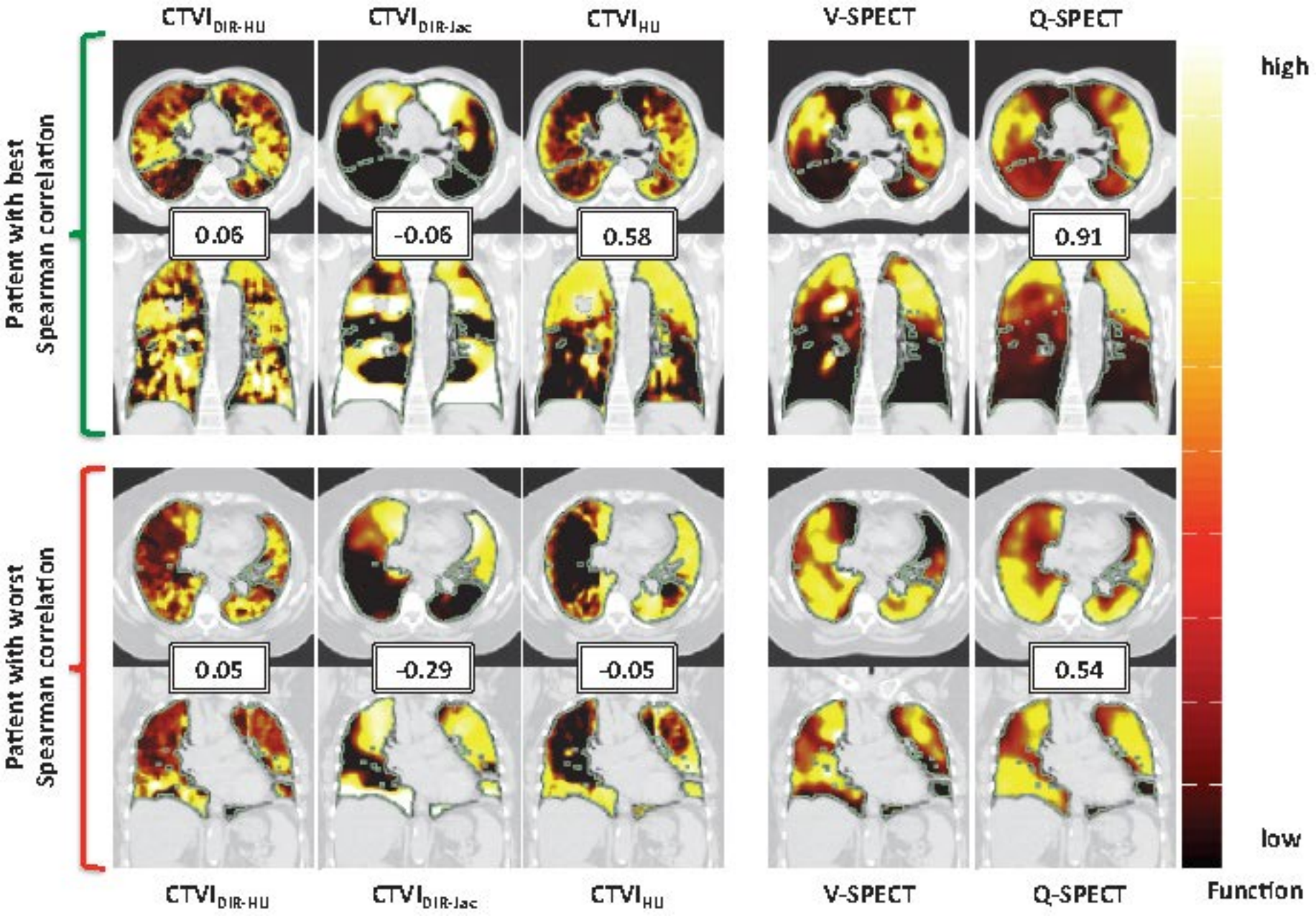
22. Vinogradskiy Y, Schubert L, Diot Q, et al. Regional lung function profiles of stage I and III lung cancer patients: an evaluation for functional avoidance radiation therapy. *Int J Radiat Oncol Biol Phys*. 2016;95(4):1273-1280.

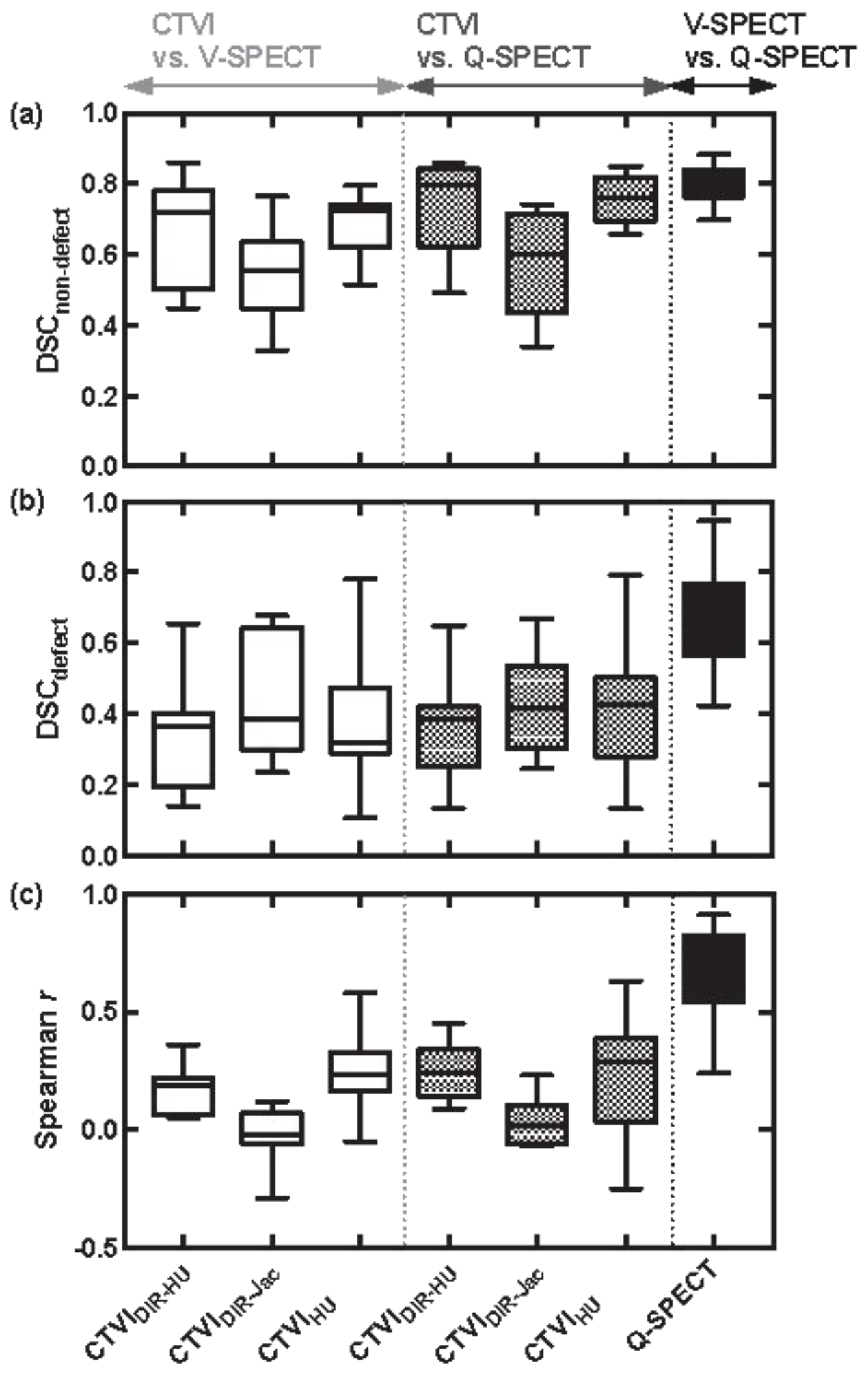
23. Yamamoto T, Langner U, Loo BW, Shen J, Keall PJ. Retrospective analysis of artifacts in four-dimensional CT images of 50 abdominal and thoracic radiotherapy patients. *Int J Radiat Oncol Biol Phys*. 2008;72(4):1250-1258.

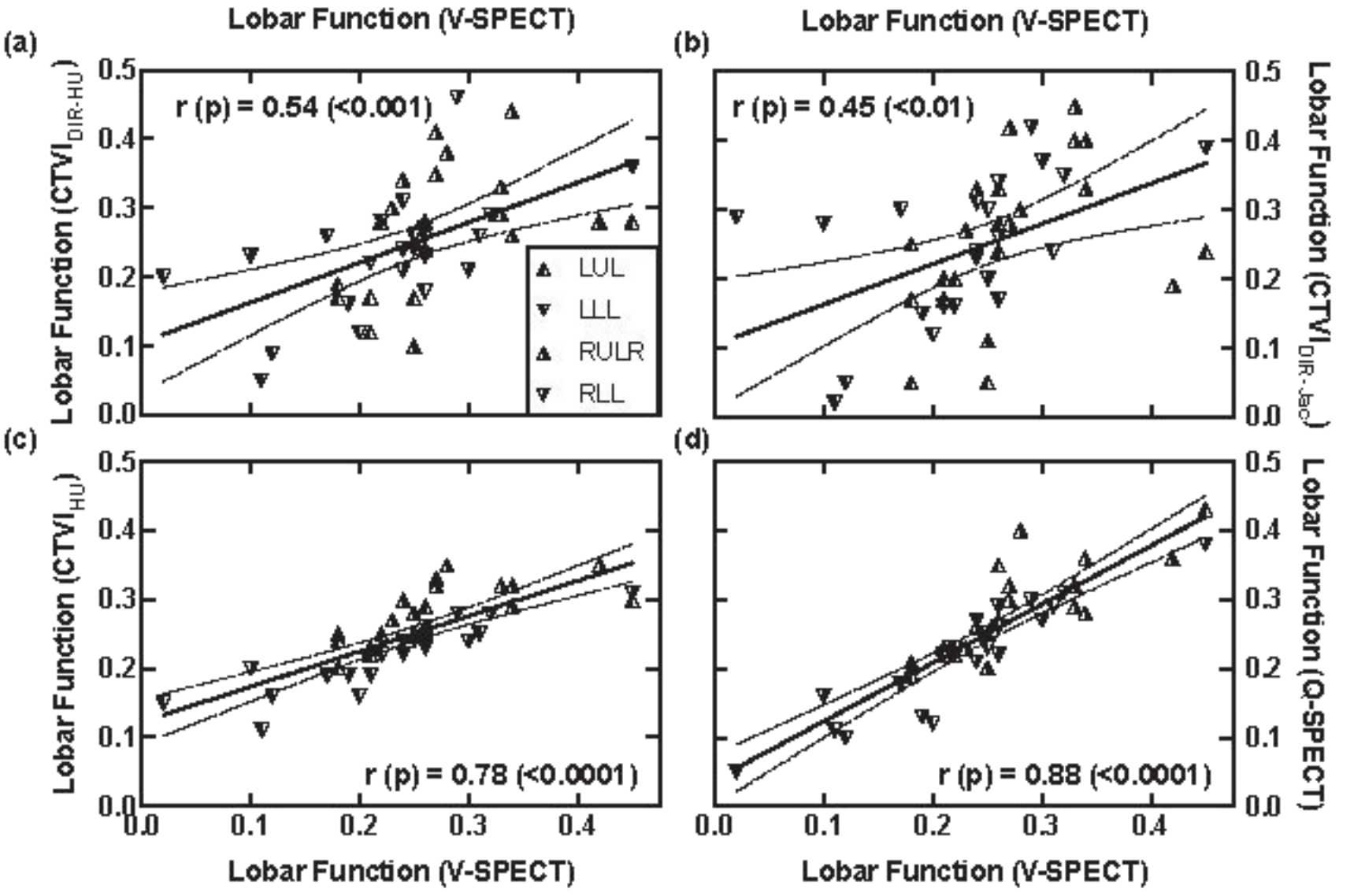
24. Yamamoto T, Kabus S, Lorenz C, et al. 4D CT lung ventilation images are affected by the 4D CT sorting method. *Med Phys*. 2013;40(10):101907.

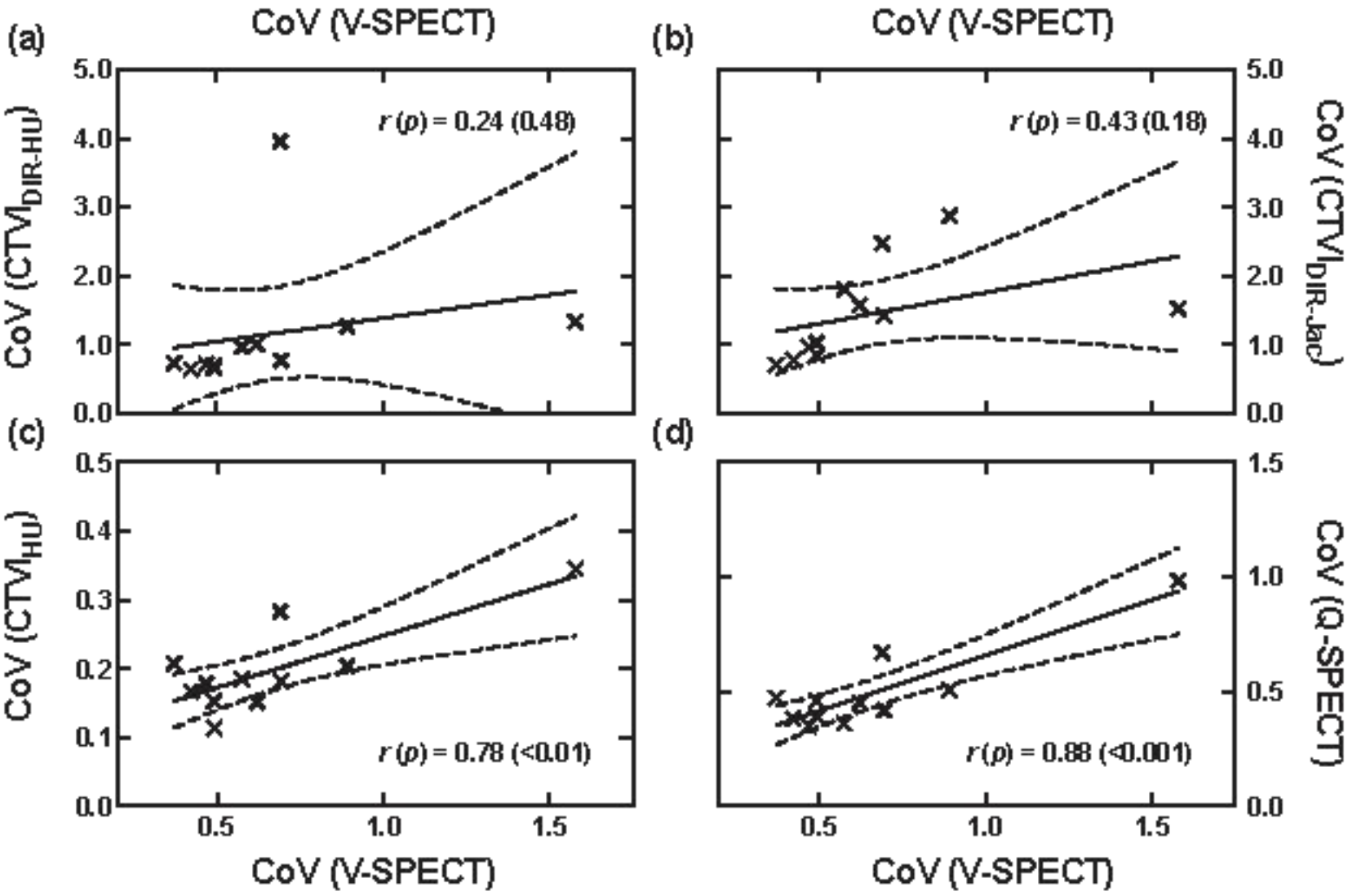
25. Cui G, Jew B, Hong JC, Johnston EW, Loo Jr BW, Maxim PG. An automated method for comparing motion artifacts in cine four-dimensional computed tomography images. *J Appl Clin Med Phys*. 2012;13(6):3638-3649.

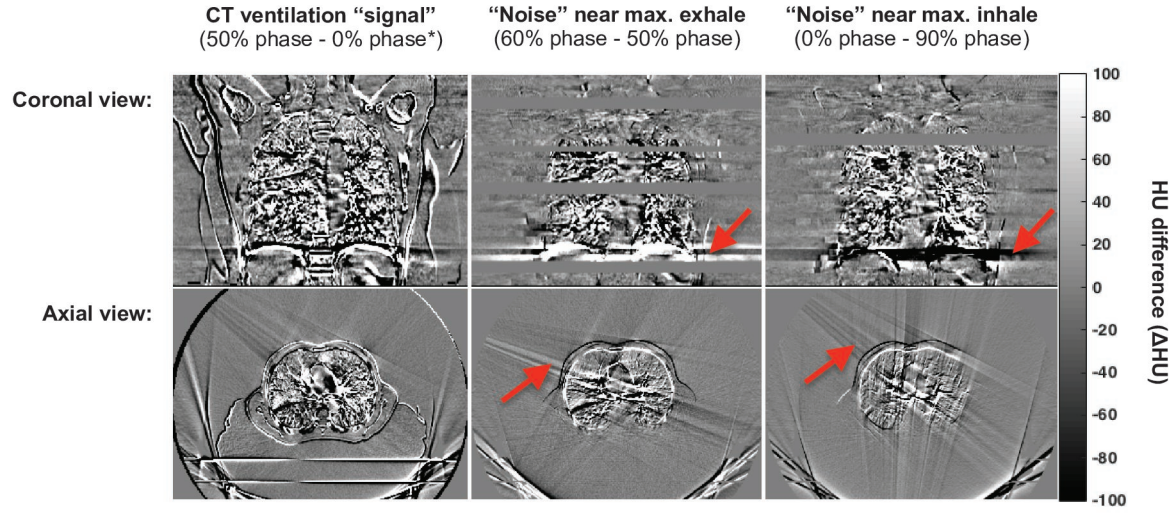
26. Latifi K, Huang T-C, Feygelman V, et al. Effects of quantum noise in 4D-CT on deformable image registration and derived ventilation data. *Phys Med Biol*. 2013;58(21):7661-7672.


CTVI DIR-HU
CTVI DIR-Jac
CTVI HU
V-SPECT
Q-SPECT
Patient with best Spearman correlation
0.06
-0.06
0.58
0.91
Patient with worst Spearman correlation
0.05
-0.29
-0.05
0.54
high
low
Function


CTVI
vs. V-SPECT
CTVI
vs. Q-SPECT
V-SPECT
vs. Q-SPECT
(a)
DSC non-defect
(b)
DSC defect
(c)
Spearman r
CTVI DIR-HU
CTVI DIR-Jac
CTVI HU
CTVI DIR-HU
CTVI DIR-Jac
CTVI HU
Q-SPECT


(a)
Lobar Function (V-SPECT)
Lobar Function (CTVI$_{DIR-HU}$)
r (p) = 0.54 (<0.001)
LUL
LLL
RULR
RLL
(b)
Lobar Function (V-SPECT)
r (p) = 0.45 (<0.01)
Lobar Function (CTVI$_{DIR-Jac}$)
(c)
Lobar Function (CTVI$_{HU}$)
r (p) = 0.78 (<0.0001)
Lobar Function (V-SPECT)
(d)
r (p) = 0.88 (<0.0001)
Lobar Function (Q-SPECT)
Lobar Function (V-SPECT)
0.0
0.1
0.2
0.3
0.4
0.5


(a)
CoV (V-SPECT)
CoV (CTV$_{DIR-HU}$)
$r(\rho)$ = 0.24 (0.48)
(b)
CoV (V-SPECT)
CoV (CTV$_{DIR-Jac}$)
$r(\rho)$ = 0.43 (0.18)
(c)
CoV (CTV$_{HU}$)
$r(\rho)$ = 0.78 (<0.01)
CoV (V-SPECT)
(d)
CoV (Q-SPECT)
$r(\rho)$ = 0.88 (<0.001)
CoV (V-SPECT)


CT ventilation "signal"
(50% phase - 0% phase*)
"Noise" near max. exhale
(60% phase - 50% phase)
"Noise" near max. inhale
(0% phase - 90% phase)
Coronal view:
Axial view:
100
80
60
40
20
0
-20
-40
-60
-80
-100
HU difference (ΔHU)