RESEARCH ARTICLE

The impact of respiratory gating on improving volume measurement of murine lung tumors in micro-CT imaging

S. J. Blocker[1], **M. D. Holbrook**[1], **Y. M. Mowery**[2], **D. C. Sullivan**[1], **C. T. Badea**[1]*

1 Center for In Vivo Microscopy, Department of Radiology, Duke University Medical Center, Durham, North Carolina, United States of America, **2** Department of Radiation Oncology, Duke Cancer Institute, Durham, North Carolina, United States of America

* cristian.badea@duke.edu

Abstract

Small animal imaging has become essential in evaluating new cancer therapies as they are translated from the preclinical to clinical domain. However, preclinical imaging faces unique challenges that emphasize the gap between mouse and man. One example is the difference in breathing patterns and breath-holding ability, which can dramatically affect tumor burden assessment in lung tissue. As part of a co-clinical trial studying immunotherapy and radio-therapy in sarcomas, we are using micro-CT of the lungs to detect and measure metastases as a metric of disease progression. To effectively utilize metastatic disease detection as a metric of progression, we have addressed the impact of respiratory gating during micro-CT acquisition on improving lung tumor detection and volume quantitation. Accuracy and precision of lung tumor measurements with and without respiratory gating were studied by performing experiments with *in vivo* images, simulations, and a pocket phantom. When performing test-retest studies *in vivo*, the variance in volume calculations was 5.9% in gated images and 15.8% in non-gated images, compared to 2.9% in post-mortem images. Sensitivity of detection was examined in images with simulated tumors, demonstrating that reliable sensitivity (true positive rate (TPR) $\geq$ 90%) was achievable down to 1.0 mm^3 lesions with respiratory gating, but was limited to $\geq$ 8.0 mm^3 in non-gated images. Finally, a clinically-inspired "pocket phantom" was used during *in vivo* mouse scanning to aid in refining and assessing the gating protocols. Application of respiratory gating techniques reduced variance of repeated volume measurements and significantly improved the accuracy of tumor volume quantitation *in vivo*.

Introduction

Preclinical imaging has become an important tool for translating imaging techniques for the detection and monitoring of cancer. A growing application of small animal imaging is its utilization in supplementing a co-clinical trial [1, 2]. However, there are critical challenges in the design and interpretation of translational animal studies that must be overcome to collect data which are meaningful in the context of clinical cancer research [3]. These include, but are not limited to, subject size, imaging hardware, motion, etc.

Competing interests: The authors have declared that no competing interests exist.

We are developing quantitative imaging methods for use in the preclinical arm of a co-clinical trial investigating synergy between immunotherapy and radiotherapy in sarcoma [4]. While magnetic resonance imaging is being employed to study primary lesion characteristics in the limbs of both patients and mice, a critical metric of therapeutic response in patients on study is metastasis-free survival determined through serial imaging of the lungs, which is the most common site for metastasis for this type of soft tissue sarcoma. To mimic the clinical study and patient standard of care, we are performing periodic micro-computed tomography (CT) scans of mouse lungs to detect and quantitate metastasis after treatment. Delineation of a significant trend in therapeutic response requires accurate and precise volumetric measurements of lung tumors, necessitating the implementation of respiratory gating techniques.

To understand the effect of respiratory gating on the sensitivity and precision of lesion detection and measurement, we performed a series of *in vivo* scans on primary lung tumor-bearing mice [5, 6]. To better understand the effect of gating on true positive and false negative rates, we also used simulated tumor images, both gated and non-gated, which could be compared to a predetermined ground truth. Additionally, we evaluated the utility of retrospectively gated images as an alternative to prospective gating with reduced scan time. Finally, we have designed a clinically-inspired "pocket phantom" (i.e., a reference standard) for external use during scanning to aid in assessing gating protocols [7]. We performed a series of *in vivo* scans on tumor-bearing mice while varying the animal's position (test-retest), and performing the analyses in triplicate (repeated measures) to assess the effects of gating on tumor detection and volumetric assessment.

Materials and methods

Animal model

All animal handling and imaging procedures were performed according to protocols approved by the Duke Institutional Animal Care and Use Committee (IACUC). A breeding pair of LSL-KrasG12D; Trp53fl/fl mice on a 129/SvJae background was originally obtained from the laboratory of Tyler Jacks (Massachusetts Institute of Technology). *LSL-Kras^{G12D}*; *Trp53$^{fl/fl}$* were maintained on a pure 129/SvJae genetic background and bred at Duke University. Animals were housed and monitored in an approved housing facility in the same building as the imaging equipment, where food, water, and environmental enrichment were provided daily. Lung tumors were generated by intranasal administration of adenovirus expressing Cre recombinase (Adeno-Cre, Gene Transfer Vector Core, University of Iowa) into *LSL-Kras^{G12D}*; *p53$^{FL/FL}$* mice as described previously [5, 6]. Mice began imaging-based screening in two-week intervals between 8–10 weeks post Adeno-Cre infection to monitor primary lung tumor development. Tumor-bearing mice (n = 13) were imaged at 12 weeks post infection. At study initiation, selected mice demonstrated multiple aggressive lung adenocarcinomas (∼0.5–1.5 mm in diameter). All animals were imaged at 24–30 weeks of age. When necessary, euthanasia was performed according to approved protocols by administering anesthetic overdose with sodium pentobarbital given intraperitoneally (IP) or intravenously (IV) at a dose of 250 mg/kg of body weight.

Micro-CT

All micro-CT imaging was performed using a micro-CT system developed in-house [8, 9]. Animals were scanned while breathing freely under anesthesia using 2–3% isoflurane delivered by nose-cone. A pneumatic pillow was positioned on the animal's thorax and connected to a pressure transducer to monitor respiratory motion and inform gating. Body temperature was maintained with heat lamps and a feedback controller. The x-ray source parameters were 80

kVp and 40 mA with 10 ms exposures. A total of 360 views were acquired over a full 360˚ rotation. The acquisition and the rotation time were on the order of 3–5 mins per set depending on the respiratory rate and the time needed for rotation to the next angle. With prospective gating, the x-ray exposure is triggered once per each respiratory cycle in the desired respiratory phase. The reconstruction resulted in a 63 μm isotropic voxel size using the Feldkamp algorithm [9] followed by bilateral filtration to reduce noise [10]. The micro-CT images were converted to Hounsfield units (HU) and saved in DICOM format. The radiation dose associated with a micro-CT scan was ~0.017 Gy per mouse, which is ~294 to 411 times less than the LD50/30 lethal dose (5–7 Gy) in mice [11].

Phantom examinations for micro-CT

To assess image quality for our micro-CT scanner, we used a Micro-CT Phantom Kit that is commercially available (http://pureimaging.com). This kit includes multiple phantoms: i) a water phantom that is used for HU calibration as well as for measuring uniformity, noise and the noise power spectrum (NPS) assessed using volumes of interest measuring 64x64x64 voxels around the periphery and at three different Z positions; ii) a calcium hydroxyapatite (HA) phantom for bone density calibration that can also be used to verify CT number linearity; iii) a phantom containing 25 μm tungsten wires designed to assess the modulation transfer function (MTF) used for determining spatial resolution; iv) a multi-layer phantom (also known as Defrise phantom) to assess the severity of cone beam artifacts when using approximate reconstruction algorithms.

Prospective gating techniques

While clinical chest CT is performed during a single breath hold, projection data in micro-CT must be acquired over many breaths. Resulting motion can lead to image blurring which confounds measurement of small structures (such as developing metastases or primary lung tumors), making respiratory gating necessary to achieve reasonable accuracy. Respiratory gating can be performed prospectively or retrospectively [12]. In prospective respiratory gating, a single respiratory phase (e.g., end expiration) can serve well to assess lung nodules. We were first to develop and implement combined cardiac and respiratory gating for micro-CT [13], which improves cardio-pulmonary imaging quality. However, the increased acquisition time which results from the addition of cardiac gating makes combined gating unfeasible for the large number of animals and frequent scan schedule required in the co-clinical trial. Consequently, we have used prospective gating to synchronize acquisition with respiration only, in which each projection is triggered when the respiratory signal breaches a user-defined threshold (Fig 1). Thus, all projections are acquired in the same phase of the respiratory cycle (e.g., end inspiration), which minimizes motion artifacts and blurring in the reconstructions [13].

Simulated tumor nodules

We have examined the effect of respiratory gating on lung tumor detection and volumetric measurement using simulated nodules inserted into temporally-resolved micro-CT datasets. A C57BL/6 mouse without lung tumors was scanned four times to acquire four different prospectively-gated respiratory phases. Gated image volumes were reconstructed using Feldkamp's algorithm (9) using 360 projections from each respiratory phase and processed with bilateral filtration [14]. A lung mask was created for each respiratory phase by thresholding the image to remove bone and soft tissue. The lung mask was dilated and eroded to make the mask internally contiguous.
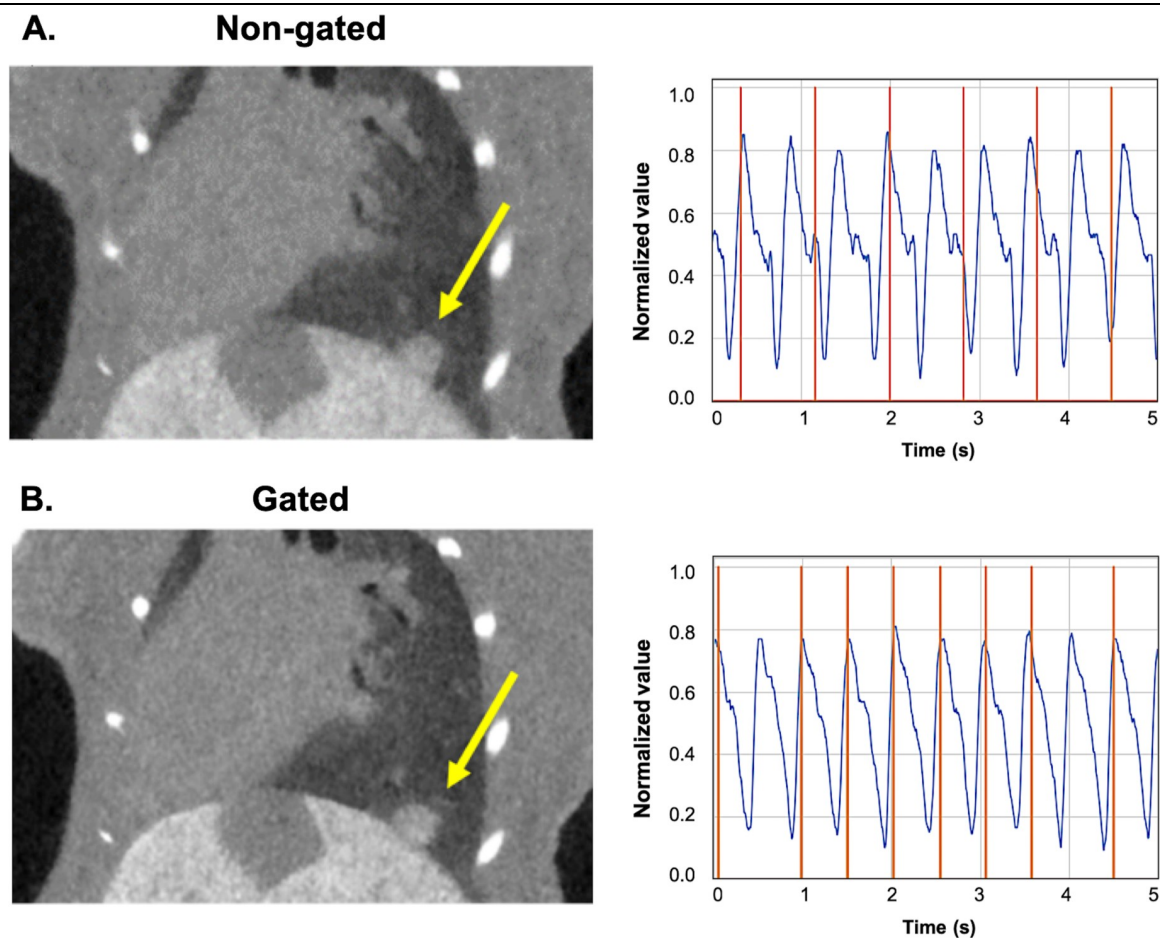
Fig 1. Prospective gating allows for synchronous image acquisition with breathing patterns. Micro-CT acquisitions without (A) and with respiratory gating (B). Shown are coronal reconstructed tomograms (left) presenting a lung lesion (yellow arrow). Normalized breath signal (blue curve) and acquisition times (orange vertical lines) are shown for each gating technique (right). Note the sharper diaphragm and tumor margins in the gated acquisition.

https://doi.org/10.1371/journal.pone.0225019.g001

Three manual segmentations of real lung tumors were taken from previous micro-CT image sets for incorporation into the reconstructions (Fig 2). To augment this small bank of tumors, they were subjected to transformations involving random rotations and skews in three dimensions, resulting in tumors with a range of sizes (0.03 to 0.28 mm^3). Transforming the tumors altered their gray-values and textures. To address this, random high frequency Gaussian noise and Gaussian filtering were applied to the tumor masks to make the augmented tumors statistically like the originals.

The augmented tumors were randomly placed within the lung mask of the first respiratory phase. The permitted locations were such that the tumor did not have more than 10% overlap outside the lung-mask or with other tumors. In real data, lung tumors are often along the edges of the lungs and near other tumors. To simulate this property, tumor locations were weighted using the inverse of Euclidian distance to the nearest lung mask boundary (i.e., positions nearest an edge of the mask will be selected with higher frequency than those far from a mask boundary). To increase the likelihood of creating groups of tumors, previously placed tumor locations were removed from the lung mask, creating boundaries with high probability for additional tumors. Once all tumors were placed in the volume, the image was forward-projected in cone
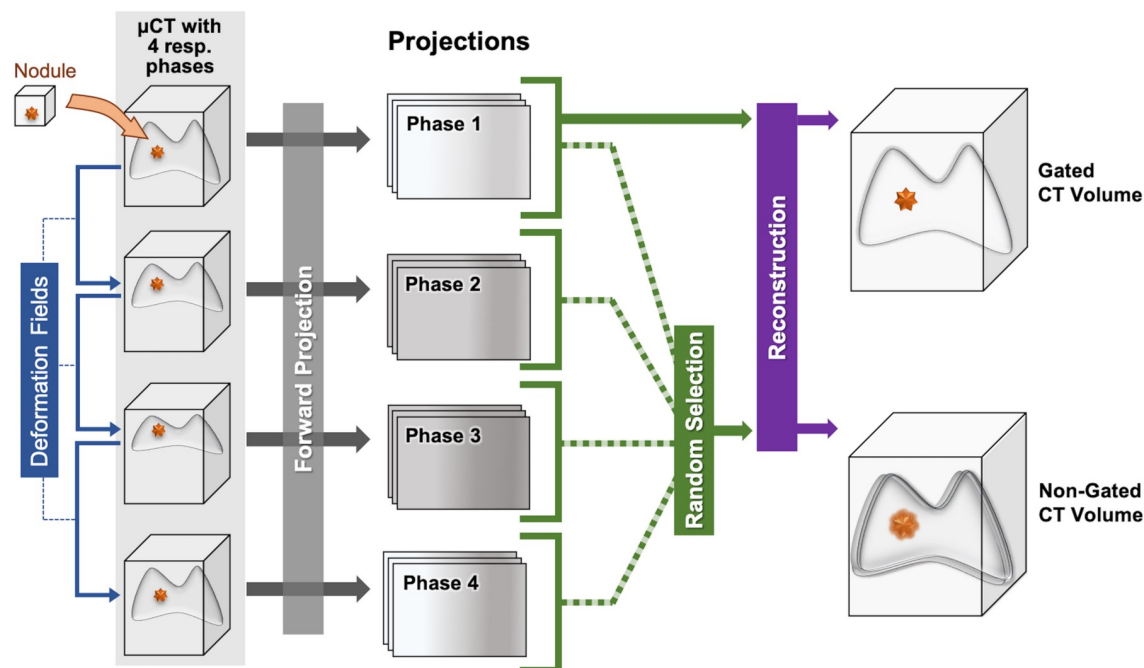
Fig 2. Process of adding simulated tumor to respiratory-gated and non-gated CT reconstructions. Deformation fields between respiratory phases are used to place tumors in volumes corresponding to different respiratory phases. Once placed, each volume associated with its lung phase is forward-projected in cone beam geometry. Random combinations of respiratory-gated projections are used to simulate non-gated acquisitions. The projections from a single lung phase create gated acquisitions.

https://doi.org/10.1371/journal.pone.0225019.g002

beam geometry to simulate CT sampling. The forward projector is part of our iterative image reconstruction GPU toolbox [15]. The object was voxelized.

To simulate non-gated acquisitions, a deformation field was computed for the lungs for each reconstructed respiratory phase using the demons algorithm [16, 17]. The deformation fields indicate how a tumor's 3D position should change with respiratory motion. The magnitudes of the computed deformation fields within the lung were small due to difficulties for the algorithm to track low-contrast features in the lungs, however the direction vectors were consistent with observed data. To correct this, we applied a multiplicative factor to the deformation fields which provided a similar magnitude of motion as in the respiratory-gated acquisitions. Tumor placement in each respiratory phase was altered from the first respiratory phase by the modified deformation field. After tumor placement, each respiratory phase image was forward-projected in the scanner cone beam geometry. Non-gated reconstruction was simulated by randomly selecting combinations of projections from each of the four respiratory phases. For gated simulations, only the first respiratory phase was used for reconstruction, mimicking the product of active breath gated acquisitions *in vivo*. Each image set contained gated and non-gated reconstructions with twelve simulated tumors.

Retrospectively gated image generation and analysis

Images of five tumor-bearing mice were acquired with prospective respiratory gating and without (non-gated). Respiration was monitored throughout the non-gated acquisition, and the resulting projections were divisible into four breath phases. Projections from a single respiratory phase (phase 0) were used as the retrospectively gated CT images for tumor segmentation. Unfortunately, retrospective gating produces an irregular angular distribution of

projections per phase; thus, when using Feldkamp's reconstruction algorithm, streaking artifacts affect the image quality of the tomographic reconstructions [18].

In a retrospective gating strategy, the projection images corresponding to a chosen respiratory phase are also fewer in number. Thus, an iterative reconstruction is needed to reduce the artifacts associated with missing angles. Reconstructions performed with fewer projections generally appear noisier than well-sampled reconstructions. To improve the reconstruction of retrospectively-gated data, we have used a sophisticated iterative reconstruction algorithm based on the split Bregman method with the add-residual-back strategy and with 4D bilateral filtration for regularization [19]. In essence, we solve the problem of temporal reconstruction (i.e., four different respiratory phases) by minimizing the following penalized, weighted least squares objective function:

$$ X = \frac{\arg \min}{X} \frac{1}{2} \sum_t [(AX_t - \mathbf{y})^T Q_t (AX_t - \mathbf{y})] + \lambda \|X\|_{\text{BTV}}. $$

Each column of the reconstructed data, X, represents the reconstructed volume at a different respiratory phase, indexed by t. The columns of X are related to the acquired projection data, $\mathbf{y}$, through the CT system projection matrix, A, and a time-point-specific, diagonal weighting matrix, Q_t. The weights on the diagonal of Q_t are repeated for each detector pixel within a single projection. The regularization term is controlled via the regularization parameter $\lambda = 0.2$. For retrospective gating, projection weights are assigned based on Gaussian basis functions which are centered on the respiratory phase to be reconstructed and which have a full-width-at-half-maximum equal to the integration time of the projection data, scaled to span the complete respiratory cycle. The iterative reconstruction was initialized with temporally-weighted least-squares evaluated for each time point, t, ($n_{t=4}$ respiratory phases) and with the biconjugate gradient stabilized convex solver (BiCGSTAB [20]). For regularization we minimize bilateral total variation (BTV) through the application of bilateral filtration (BF, [21]). The filtration domain is expanded to include the respiratory phase point proceeding (t-1) and following (t+1) the time point being filtered.

Average signal-to-noise (SNR) was measured in regions of interest (ROIs) drawn in the liver and empty space for comparison between prospectively gated, retrospectively gated, and non-gated images. Three tumors from each mouse were selected in the prospectively gated image, varying the size and relative location as much as possible. Tumor isocenter coordinates were recorded for consistent localization across segmentations. Measurements were performed in triplicate in each image type, alternating between prospective, retrospective, and non-gated. Sets within each triplicate were performed with at least one day in between measurements, to avoid serial-segmentation biases in edge detection. Resulting tumor volumes were compared within each individual image (test-retest of measurements to determine measurement variance), as well as between image types (to determine accuracy and reproducibility under different gating conditions).

Phantom studies

Previously, we have used x-ray radiopaque beads surgically glued to the abdominal diaphragmatic surface and cardiac ventricles of anesthetized, mechanically ventilated rats to describe the range of motion using digital microradiography with respiratory gating methods [22]. For this study, a mouse phantom similar to the CT human "pocket phantom" [5] was fabricated using a number of polyethylene spheres (http://www.cospheric.com/) with diameters of ~0.7 mm. The spheres were embedded in a packaging foam, which imitates the appearance of lung parenchyma, and loaded into a microcentrifuge tube. The complete phantom was positioned

on the animal chest to freely move with respiration, and analysis of the beads was used as a reference to quantify respiratory gating performance (Fig 3).

Mouse scans

For preliminary *in vivo* lung tumor studies, we scanned the same mouse in three adjusted positions (repositioned on scan bed), acquiring images with and without gating in each position. A non-gated scan requires about approximately 3 minutes and a prospectively gated scan takes about 5 minutes to acquire. Post-mortem *ex vivo* scans of the same mouse were also performed in triplicate and used as the standard for defining tumor volumes and segmentation precision.

Subsequent in vivo imaging was performed by acquiring one gated and one non-gated image (as described above) in a single position. All acquisitions were step-and-shoot and with a circular cone beam trajectory. The field of view was large enough to cover the whole mouse since the sensitive area of the detector is 145.4 mm x 114.9 mm.

Image analysis

All measurements, including segmentations and SNR measurements, were performed in triplicate. A collection of identifiable tumors *in vivo*, phantom pellets, or simulated tumors were identified in each of the scan sets. Volumes were calculated via manual segmentation in OsiriX DICOM viewer or 3D Slicer (https://www.slicer.org/) image analysis software. For simulated images, the reader was blinded and asked to identify as many tumors as possible, as well as to segment each lesion to calculate a volume. Subsequently, the reader performed manual segmentations on the binary label map containing all simulated tumors ("truth") to measure the precision of the segmentation method. Each tumor identified in the gated and non-gated images (including true positives and false positives) was compared to the binary label map to calculate detection sensitivity and precision, as well as volumetric accuracy. *In vivo*, tumors and phantom pellets were identified for each condition (gating, no gating, post-mortem) and each repositioning on the cradle, resulting in nine measurements per volume of interest.

Establishment of a quantitative imaging biomarkers alliance (QIBA)-inspired profile for volumetry

In order to understand the effect of precision on day-to-day tumor assessments for our co-clinical trial, we have implemented a QIBA-inspired profile for mouse lung tumor volumetry in CT with gating techniques based on tumor size [23, 24]. In a pilot group of tumor-bearing mice, images were acquired with and without prospective respiratory gating. After scanning, the animals were sacrificed and post-mortem images were acquired. Lung nodules (n = 13) were identified and measured in triplicate in each of the 3 conditions. Coefficients of variance (CV) were calculated for each measured tumor volume in gated, non-gated, and post-mortem conditions. The tumors were then divided into 3 groups based on size (as measured post-mortem). Within-subject coefficients of variance (wCV) were determined for each size group in gated and non-gated images [25], and used to calculate the 95% confidence intervals (CI) as follows:

$$95\% \, CI = Y \pm (1.96 * Y * wCV)$$

where Y is the measured tumor volume, and wCV is dependent upon the size of Y.

To assess the utility of the calculated 95% CIs, an independent group of tumor-bearing animals were scanned with and without respiratory gating. Three tumors of varying sizes which could be reliably identified in both gating conditions were selected in each mouse (n = 15),
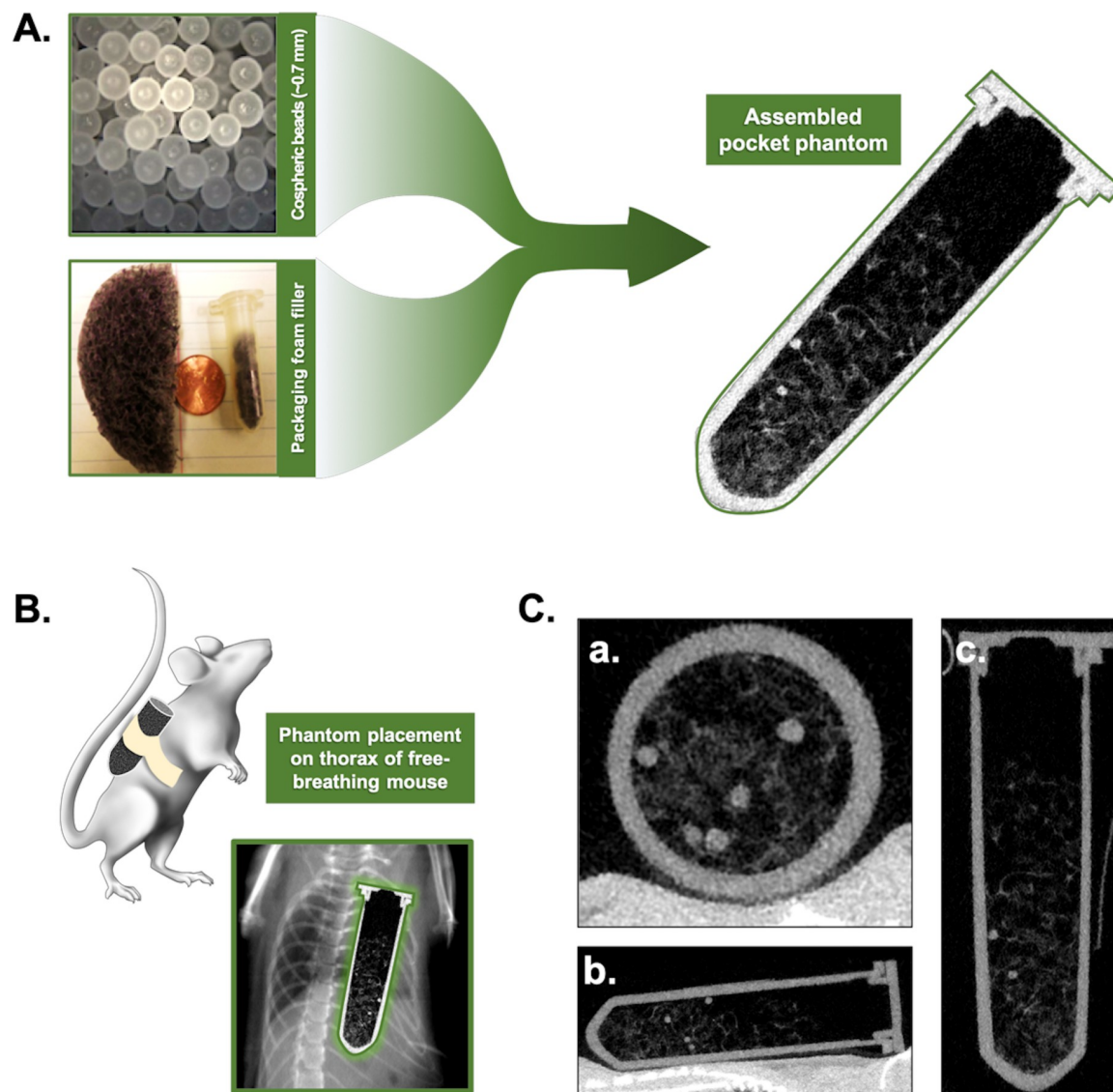
Fig 3. Design of a pocket phantom for evaluation of respiratory gating. The phantom was constructed using packaging foam to mimic the lung parenchyma in which a number of polyethylene spheres were inserted (A). The constructed pocket phantom was fixed to the thorax of free-breathing mice, allowing for uninhibited motion throughout the respiratory cycle (B). Examples of the micro-CT images of the pocket phantom (C) are shown in the axial (a.), coronal (b.), and sagittal (c.) view.

and volume measurements were repeated three times for each tumor/condition. The results were plotted against the calculated 95% CIs to determine if the size-dependent prediction of precision reasonably reflected variance of repeated measures across a range of tumor sizes.

Statistical analysis

Analysis of variance (ANOVA) calculations were performed to examine the effect of gating on volume reproducibility. Bland-Altman analyses were used when comparing acquisition techniques in simulated and retrospectively gated images. Statistical significance was defined as: $^*p<0.05$, and confidence intervals were calculated at 95%. Statistical analyses were performed

and visualized using GraphPad Prism version 7.00 for Mac (GraphPad Software, La Jolla California, US).

Data access

Access to image files utilized during this study have been made publicly available via the Duke Voxport/Voxstation [26]. Access to our Voxport Image Management System is available at: https://civmvoxport.vm.duke.edu/voxbase/index.php. The visitor will be prompted to login to CIVMVoxPort using the provided credentials (User Name: DukeU24, Password: DukeU24Review). Data relevant to this study will appear under the folder labeled "Micro-CT with Respiratory Gating".

Results

Technical performance of the micro-CT scanner was assessed with phantom scans (Fig 4). The mean noise values (standard deviation) in the water phantom were around 68 HU (Fig 4A). Note that the levels of noise are much higher than in clinical CT due to the higher resolution. However, these noise levels are acceptable when imaging a high contrast organ such the lungs and using very low radiation dose as is required for longitudinal imaging. The line profile shows good uniformity with no cupping. The system has linear response (Fig 4B) and spatial resolution given by the MTF at 10% is 5.5 line pairs/mm (Fig 4C). The NPS measured in the water phantom indicates some structured noise. The noise power decreases significantly at high spatial frequencies (Fig 4D). We also show that with our scanner, the cone beam artifacts are acceptable when using Feldkamp algorithm with a circular scanning trajectory (Fig 4E). Our cone beam geometry creates cone angle of ~10 degrees.

Respiratory gating reduces variance and improves accuracy of tumor volume measurements *in vivo*

Under anesthesia during imaging, mouse respiratory rates were typically maintained between 80–120 breaths per minute. To quantify the impact of respiratory gating on tumor volume measurements, selected lung lesions in mice with primary tumors were segmented in triplicate from images acquired with gating, without gating, and after sacrifice (motion-free control). Tumors of varying size which were locatable in all three acquisition conditions were selected for analysis (Fig 5). The mean volume calculated in non-gated images was significantly different (*p<0.05, calculated via ANOVA) from the post-mortem values in 2 out of 4 tumors (Fig 5B and 5C). Gated volumes did not differ significantly, suggesting that gating improves accuracy of volume measurements compared to non-gated images. In all cases, the mean volume calculated from the non-gated image was greater than both the post-mortem and gated images. On average, the variance in volume calculations was 2.9% in post-mortem images, 5.9% in gated images, and 15.8% in non-gated images (Fig 5D). Thus, prospective gating reduces variance compared to non-gated images in repeated measures, demonstrating improved volume measurement precision. Gating did not result in higher variance or lower accuracy than non-gated images when calculating tumor volumes in any case. Taken together, these data suggest that in free-breathing mice, prospective gating reduces motion-based blurring and improves tumor volumetry precision and accuracy.

From data acquired in animals scanned before and after sacrifice, we constructed a QIBA-inspired profile for lung tumor segmentation in CT images acquired with and without gating on our scanner. Given the dependency of segmentation precision or accuracy on nodule size, the predicted confidence intervals were calculated for ranges of tumor volumes observed on study. As expected, wCV and confidence intervals were generally reduced for volume
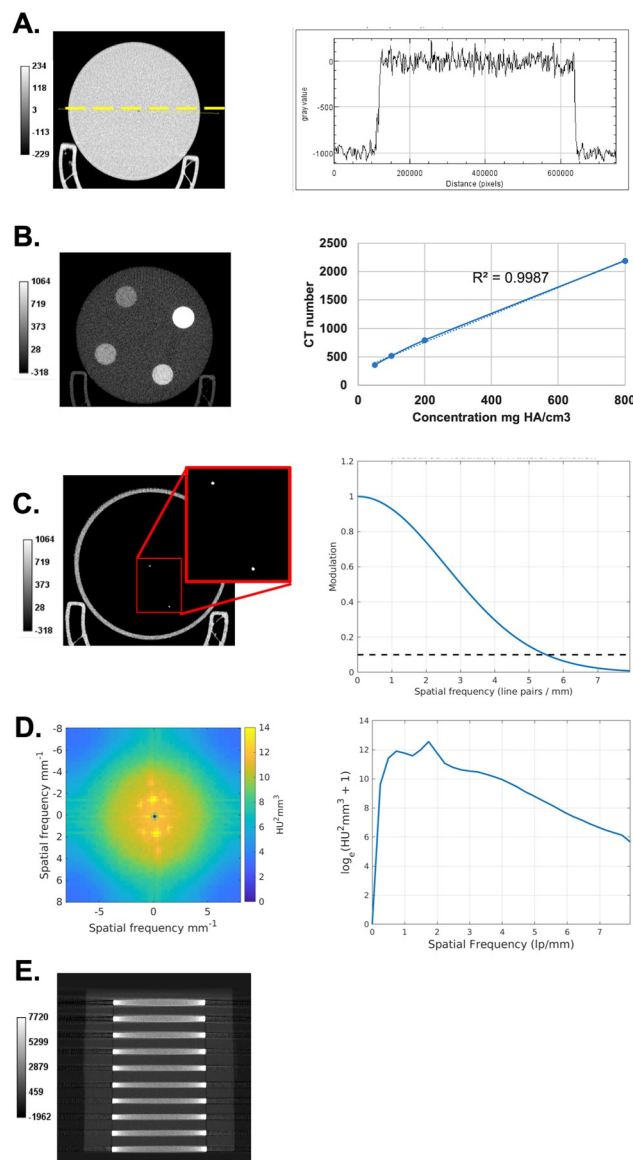
Fig 4. Phantom assessment of the micro-CT scanner performance. Shown is a slice and a line profile through the water phantom (A). The CT linearity was obtained by scanning a phantom containing 4 inserts with various concentrations of calcium hydroxyapatite (B). The spatial resolution was assessed with a wire phantom (C). Shown also is the image and the radial profile of NPS obtained in the water phantom (D). Finally, a coronal slice and a line profile of the multi-layer Defrise phantom is shown (E). No substantial cone beam artifacts are present.

https://doi.org/10.1371/journal.pone.0225019.g004

assessments in gated images compared to non-gated images (Fig 6A). This disparity increased with rising nodule volume, where the precision of gated volumetry improved while non-gated precision remained unchanged. To assess the utility of the derived confidence intervals for gated and non-gated volumetry, we performed test-retest triplicate segmentations of multiple tumors in five mice. In gated and non-gated images, the calculated 95% confidence intervals were successful in predicting variance of volume measurements based on tumor size, with only one notable exception in a non-gated segmentation of a large tumor (Fig 6B).
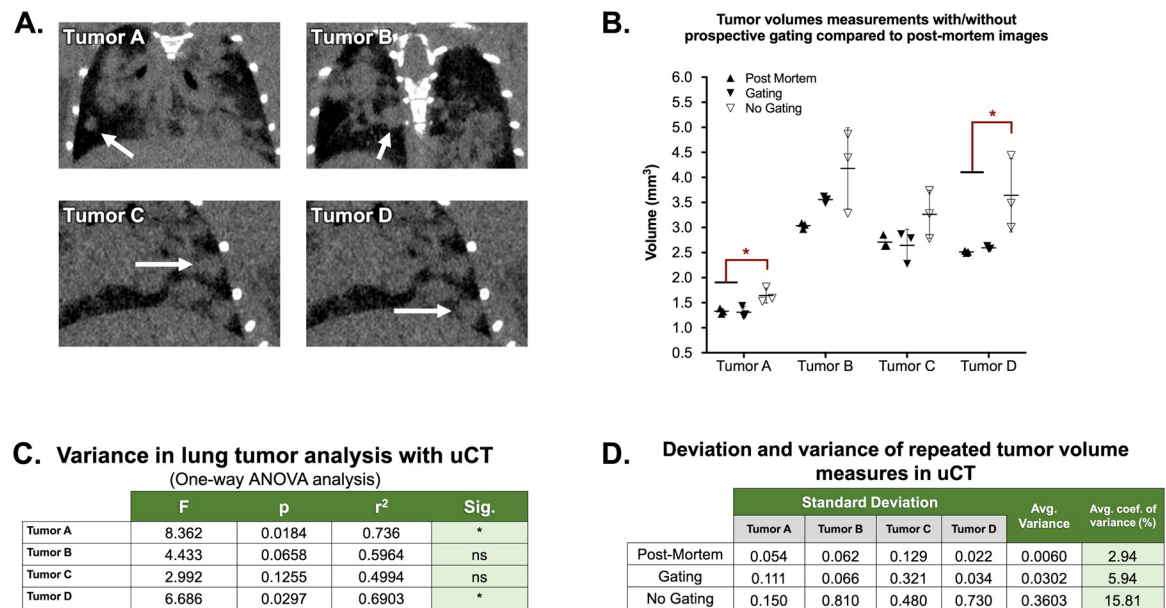
A.

Tumor A Tumor B

Tumor C Tumor D

B.

Tumor volumes measurements with/without prospective gating compared to post-mortem images

C. Variance in lung tumor analysis with uCT
(One-way ANOVA analysis)

	F	p	r²	Sig.
Tumor A	8.362	0.0184	0.736	*
Tumor B	4.433	0.0658	0.5964	ns
Tumor C	2.992	0.1255	0.4994	ns
Tumor D	6.686	0.0297	0.6903	*

D. Deviation and variance of repeated tumor volume measures in uCT

	Standard Deviation				Avg. Variance	Avg. coef. of variance (%)
	Tumor A	Tumor B	Tumor C	Tumor D		
Post-Mortem	0.054	0.062	0.129	0.022	0.0060	2.94
Gating	0.111	0.066	0.321	0.034	0.0302	5.94
No Gating	0.150	0.810	0.480	0.730	0.3603	15.81

Fig 5. Respiratory gating improved the accuracy and precision of tumor volume measurements in the lung. Volumes were calculated for a collection of tumors (A) in 3 positions per condition, including a non-gated image, a gated-image, and an image acquired post-mortem (standard). When accounting for variance, volumes from non-gated images were significantly different from post-mortem standards in half of cases, while gated image volumes did not differ significantly (B; *P<0.05, ANOVA). The resulting ANOVA table describes the deviation and difference in means between acquisition condition per tumor (C). Coefficients of variation were calculated for each condition and averaged for each acquisition type (D).

https://doi.org/10.1371/journal.pone.0225019.g005

Simulated lung nodules clarify the sensitivity and specificity of tumor detection in micro-CT of mouse lungs with and without respiratory gating

To assess false negative rates and limits of detection, we utilized our lung micro-CT datasets containing simulated tumors which can be compared to a binary label map, or absolute truth (Fig 7). Simulated tumor volumes were measured in triplicate on the original binary label map (to quantitate any variance due to segmentation drawing), the non-gated images, and the gated images (Fig 8A). Precision of repeated volume measurements was frequently superior in the gated acquisitions compared to the non-gated measurements. Additionally, the accuracy of tumor volume calculations in gated images was consistently better than in non-gated images, with multiple non-gated volume calculations deviating significantly from the binary label map (truth).

Bland-Altman analysis comparing volume measurements from binary label maps with those derived from gated and non-gated acquisitions suggested that both acquisition types resulted in volume measurements of reasonable agreement with the ground truth (Fig 8B). However, the trend seen in the plot of non-gated acquisitions, demonstrated by a linear negative slope, suggests that volume calculations from these images are subject to a tumor size-dependent bias. Specifically in non-gated images, large simulated tumors and small simulated tumors were frequently over- and under-estimated, respectively.

Finally, with the availability of ground truth we were able to identify the limitations in sensitivity of both gated and non-gated acquisitions. The occurrence of both false negative and false positive lesion identification was greater in images without respiratory gating. The true positive rate (TPR; sensitivity) in correctly identifying tumors was improved with the application of gating to 0.97, compared to 0.56 in non-gated images (Fig 8C). Similarly, the positive predictive value (PPV; precision) of gated images was 1.0, compared to 0.71 in non-gated images.
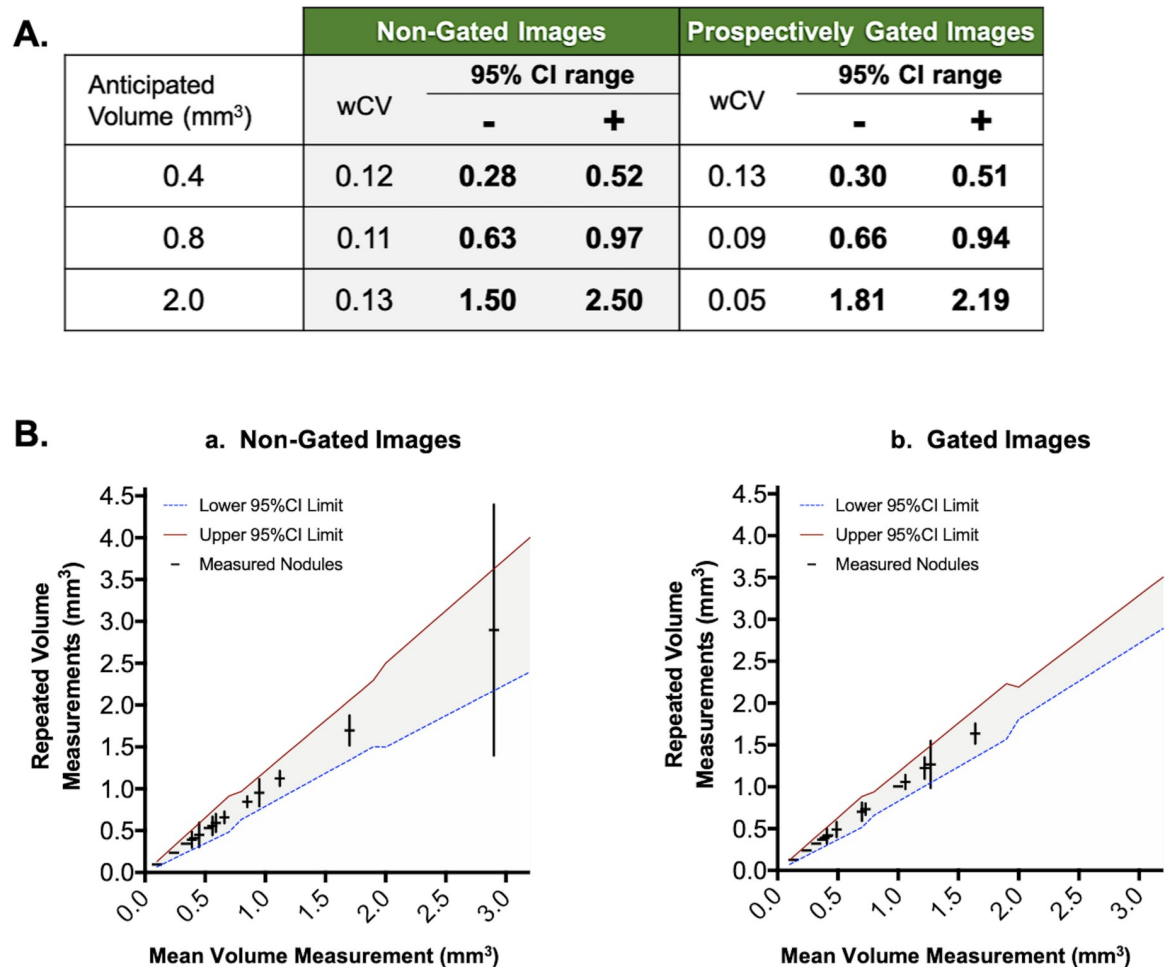
A.

Anticipated Volume (mm³)	Non-Gated Images			Prospectively Gated Images		
	wCV	95% CI range		wCV	95% CI range	
		-	+		-	+
0.4	0.12	**0.28**	**0.52**	0.13	**0.30**	**0.51**
0.8	0.11	**0.63**	**0.97**	0.09	**0.66**	**0.94**
2.0	0.13	**1.50**	**2.50**	0.05	**1.81**	**2.19**

B.

Fig 6. QIBA-inspired profile for predicting tumor volumetry precision across variable tumor sizes in gated and non-gated images. Shown is a table of sample anticipated tumor volumes for which the wCV and upper/lower 95% CIs was determined in both non-gated and gated acquisitions (A). Triplicate tumor measurements (plotted as mean ± standard deviation) from an independent cohort of mice were plotted against the calculated 95% CIs (red and blue bounding lines) (B), demonstrating the validity of the size-dependent confidence prediction for both non-gated (a) and gated (b) images.

https://doi.org/10.1371/journal.pone.0225019.g006

When comparing limits of detection of TPR = 0.9 (the size of lesion at which TPR drops below 0.9), gating maintained reliable detection down to tumor volumes of approximately 1.0 mm³, while non-gated images were not reliable in positive prediction at volumes below approximately 8.0 mm³ (Fig 8D).

Retrospective gating techniques improve volume measurement accuracy compared to non-gated images

Datasets including prospectively gated and non-gated images were acquired in five mice. Retrospective images were generated from the non-gated images by selecting projections from a single phase of respiration. First, signal-to-noise (SNR) measurements were taken by selecting ROIs in the liver and empty air surrounding the mouse (Fig 9A). Prospectively gated and non-gated images demonstrated similar SNRs, with prospective gating being slightly better. Retrospectively gated images showed SNR of approximately half of prospectively gated images (Fig 9B). This was expected, given that retrospectively gated images although reconstructed using
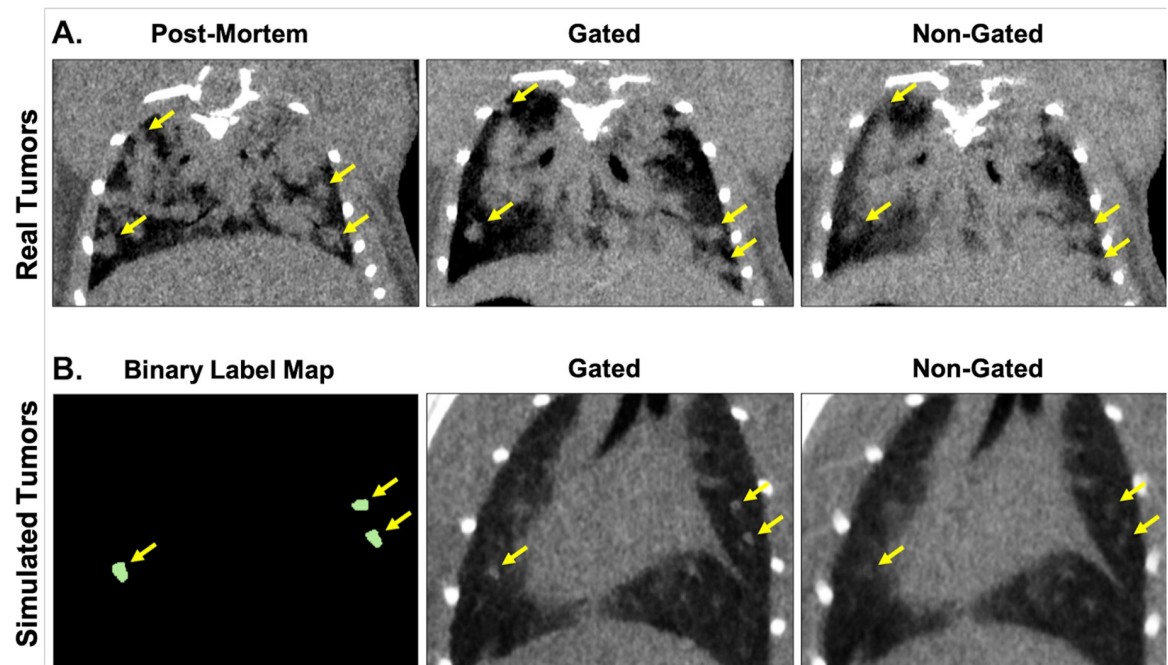
Fig 7. Demonstration of simulated lung tumor images compared to in vivo images of tumor-bearing lungs. *In vivo* lung images from a tumor-bearing mouse demonstrate the difference between non-gated and gated images compared to post-mortem tumor detection (A). Simulated tumors are created from segmentations of real tumors, which are projected into the lungs of a healthy, free-breathing mouse (B). Simulated tumors mimic the appearance and location of tumors seen within *in vivo* images, as well as the effects of respiratory gating on tumor edge definition, but have the advantage of a binary label map which serves as "truth". Real and simulated tumors are indicated with yellow arrows on *in vivo* and simulated acquisitions, respectively.

https://doi.org/10.1371/journal.pone.0225019.g007

an iterative approach use mostly one fourth of the data, resulting in more noise. Nevertheless, we note that iterative reconstruction does help compared to filtered back projection for the same amount of data.

Three tumors from each mouse were measured in triplicate in all three gating conditions, with each measurement occurring at least one day apart to avoid bias. Based on ANOVA comparisons, there were four tumors in which the non-gated image resulted in significantly different volume measurements compared to prospectively gated images (Fig 9C). Retrospectively gated images differed significantly from prospective gating in the case of two tumor volume measurements, one of which was also mis-measured in non-gated images. Bland-Altman comparison of both alternatives to prospective gating demonstrated that retrospective gating resulted in better agreement, with a smaller bias and reduced limits of agreement (Fig 9D and 9E). Overall, retrospective gating improved volume accuracy relative to non-gated images, but was not a perfect substitution for prospective gating. This is likely due to the variance in measurements caused by a lower SNR. Taken together, these data suggest that retrospective gating helps alleviate some, but not all, of the motion-related issues with tumor volumetry with a shorter acquisition time than prospectively gated lung CT.

Use of a pocket phantom demonstrates respiratory motion in real time

Pellets contained in a pocket phantom on the thorax of tumor-bearing mice were used to assess the effect of prospective respiratory gating in real time during *in vivo* imaging. The variance of pellet volume measurements was determined from three repetitions in each of the following conditions: no gating, respiratory gating, post-mortem (motion-free) (Fig 10). Three identifiable pellets were selected to mimic conditions that may be encountered during lung
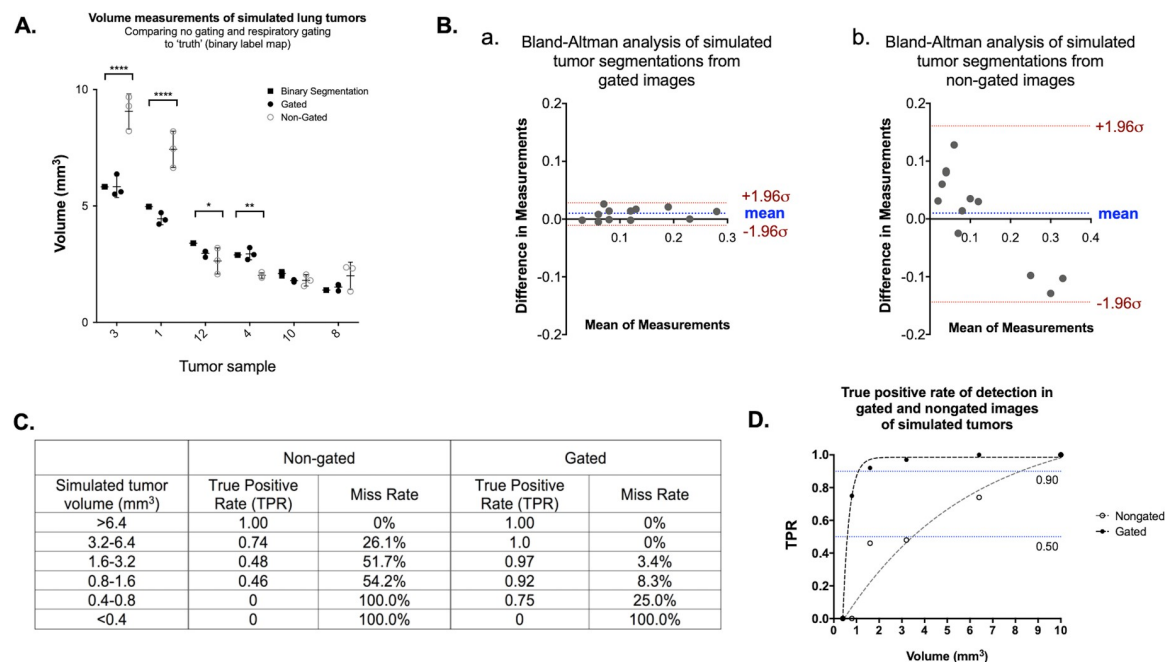
Fig 8. Simulated tumors in CT of free-breathing mouse lungs demonstrate enhanced sensitivity, precision, and volume accuracy when identifying tumors in gated images compared to non-gated images. Variably sized samples of measured simulated lung lesions are shown, including triplicate measures of binary label maps (truth), gated images, and non-gated images (A; ANOVA, *p<0.05). Bland-Altman plots of gated (B.a.) and non-gated (B.b.) tumor segmentations suggest agreement with segmentations of the binary label map, with a size-dependent bias noted in non-gated image segmentations. True positive rate (TPR; sensitivity), Miss rate (False negative rate, or lesions which the reader was unable to correctly identify), Positive predictive value (precision), False discovery rate (False positive rate), and F1 (harmonic mean of precision and accuracy) are shown for gated and non-gated images of simulated tumors of varied sizes (C). TPR as a function of simulated lesion size, with TPRs of 0.9 and 0.5 indicated (blue lines) (D).

https://doi.org/10.1371/journal.pone.0225019.g008

tumor measurement, including a solitary pellet, a pellet adjacent to the tube wall, and a pellet directly adjacent to another pellet.

Discussion

The goal of this work was to assess the performance and utility of respiratory-gated CT for detection and measurement of tumors in mouse lungs to serve a co-clinical imaging study. When compared to clinical CT, the first challenge in implementing preclinical imaging is ensuring adequate and comparable instrumentation. The performance of the micro-CT scanner was found to be adequate for imaging of small tumor nodules in the lungs of mice. We note that there are important constraints on radiation dose levels which are intentionally very low to minimize any influence of longitudinal scanning on the outcome of the experiments.

Unlike clinical chest CT, in which patients may hold their breath to reduce motion during acquisition, breathing motion in micro-CT of mouse lungs is a persistent barrier to small lesion detection. The development of lung metastases is a critical metric for assessing therapeutic response in our co-clinical trial of sarcoma, making the ability to discern differences in small lesions between treatment arms paramount. Thus, we employed respiratory gating techniques to lessen the effect of motion in the lungs, increase visibility of small lesions, and improve tumor volume measurement. *In vivo* studies demonstrated that tumor segmentations performed in images with prospective respiratory gating more accurately measured the volume of tumors than non-gated segmentations, compared to post-mortem controls. In multiple instances, non-gated tumor volumes were significantly different than control measurements,
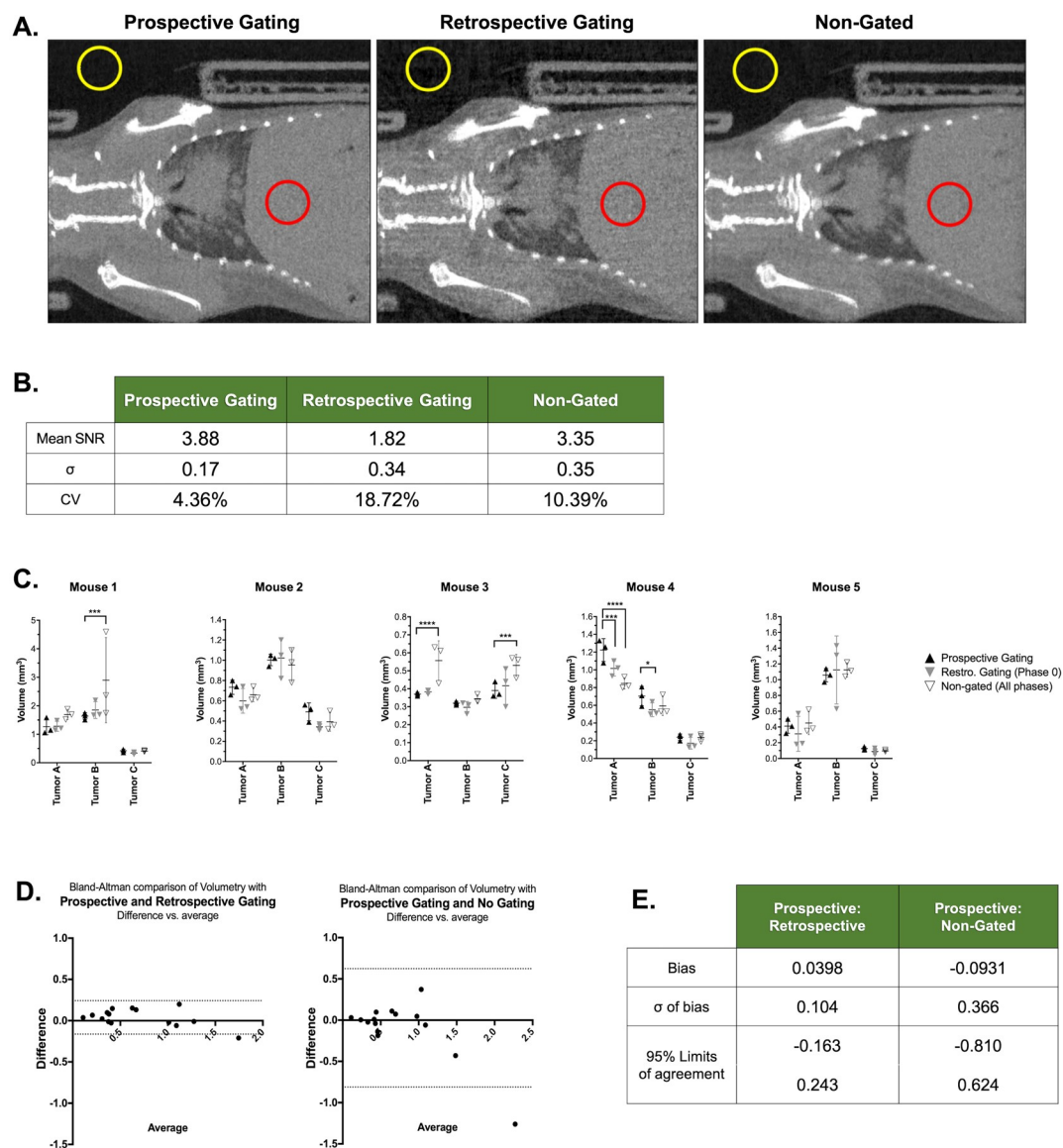
Fig 9. Retrospective gating improves tumor volume accuracy compared to non-gated images. Representative images of prospectively gated, retrospectively gated, and non-gated images were acquired in a tumor-bearing mouse (A). SNR measurements were calculated from liver (red) and empty space (yellow) ROIs to determine mean SNR and variance (B). Three tumors per mouse (n = 5) were selected and segmented in triplicate under each scan condition, and results were compared with ANOVA (C). Bland-Altman comparisons to prospective gating were performed for retrospective gated and non-gated tumor analyses (D & E).

while gated volumes never differed significantly. Further, gated segmentations were consistently more precise than those measured in non-gated images. The derived 95% confidence intervals provided a reasonable metric for estimating the probable variance in tumor volume measurements acquired with our lung CT protocols. This is of immense value in a large-scale, co-clinical trial setting, where high animal numbers make persistent measures of precision infeasible. In addition, the need for longitudinal assessments requires a means of understanding the differences in precision as a tumor changes size, without the validation of post-mortem ground truth. Overall, these data support the use of prospective respiratory gating to improve tumor analysis in the lungs of free-breathing mice.
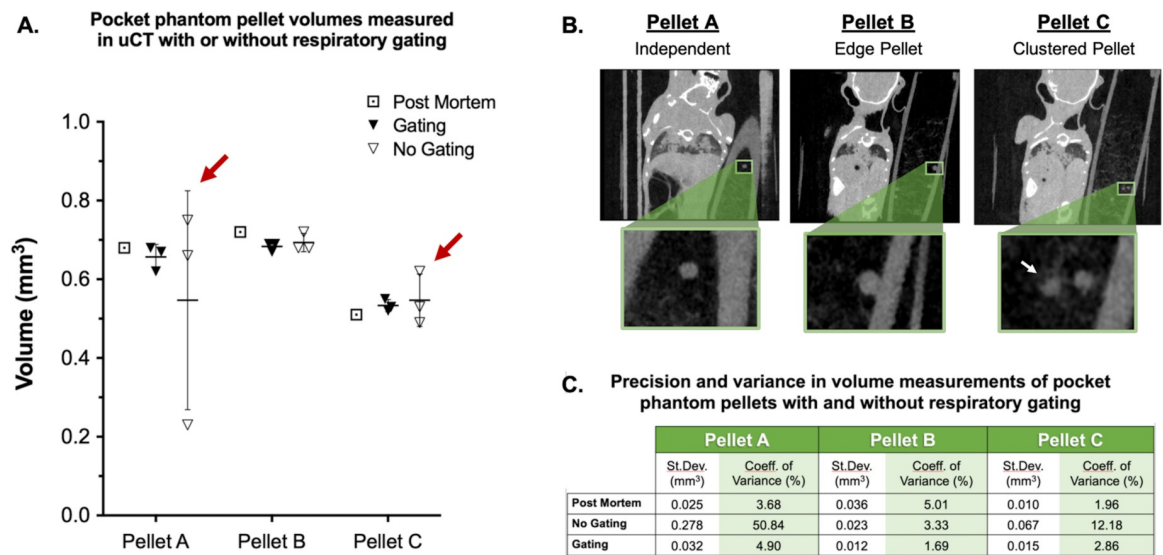
Fig 10. The effect of gating on volumetric output and measurement variance. Three pellets were measured three times in each acquisition condition, and volumes were calculated and compared to post-mortem measurements (A). Pellets were selected to mimic conditions regularly encountered during lung tumor analyses (B). Variance among measured volumes was greatest in non-gated images (red arrows in A), while gating resulted in precision similar to post-mortem measurements (C).

While post-mortem tumor measurement can serve as a motion-free control, the volumetric limits of detection with or without gating remain undefined. Specifically, post-mortem controls do not provide a complete assessment of false negative and false positive rates among image readers, even though motion is not present. To delineate the effects of respiratory gating techniques on reader interpretation, we utilized images which contained simulated tumors. These artificial tumors are designed to visually mimic real tumors, and are projected into healthy, free-breathing micro-CT images of the mouse lung. Most importantly, the size, shape, and location of simulated tumors which appear in the final projections are known, providing a definitive ground truth to which reader analyses can be compared. Using this system, we were able to interrogate the sensitivity and precision of tumor identification by readers studying our micro-CT images, as well as quantitate the advantage of respiratory gating techniques with respect to lesion size. In fact, respiratory gating improved overall sensitivity and specificity of tumor identification compared to non-gated projections. Further, accuracy of tumor volume measurement was consistently better in gated images, while non-gated images demonstrated size-dependent discrepancies in volumetric accuracy. Thus, simulated tumor images provided foundational information about how gating affects reader interpretation of preclinical lung CT, including improvements in volumetry accuracy, and size-based limits of lesion identification. Additional studies with alternative measurement techniques, such as 3D histological assessment of tumor-bearing lungs, could additionally enhance our understanding of edge detection/interpretation with and without respiratory gating. In particular, these studies would address limitations of lesion detection beyond motion, such as tumor transition regions, and differences in soft tissue composition.

While prospective respiratory gating significantly improved tumor detection and measurement accuracy/precision, there is a time cost for acquiring prospectively-gated images. Thus, we evaluated retrospective gating of non-gated images as an alternative to prospective gating. Conceptually, retrospective gating would reduce motion-related artifacts and blurring, improving lesion detection/measurement with the same short acquisition time of a non-gated image. We found that, indeed, retrospective gating did improve accuracy of tumor volumetry

compared to non-gated images. However, retrospectively-gated images were consistently noisier, resulting in less precise measurements than both non-gated and prospectively-gated images. Thus, the utility of retrospective gating for tumor volume measurement lies somewhere between acquisitions without gating and prospectively gating. It is also worth noting that while some time is saved during acquisition compared to prospective gating, long reconstruction times ($>$ 30 mins) associated with iterative reconstruction algorithms must be considered. Still, retrospective gating could serve as a reasonable alternative to prospective gating for those hoping to improve non-gated lung imaging without sacrificing scan time. We note that with more sophisticated reconstruction iterative algorithms, the performance of the retrospectively gated images can be improved [27].

Finally, given the impact of successful gating on image quality, we sought to create and implement a "pocket phantom" with which we could assess gating success on a scan-to-scan basis. We measured beads contained within the pocket phantom, which moved with the mouse as it breathed, with the same methods used to measure tumors. As was seen with lung lesions, gating improved precision when repeatedly measuring the volume of multiple pellets. The use of these pocket phantoms represents a simple and elegant technique by which users can monitor gating success *in vivo*. Specifically, these phantoms may be used as internal controls to identify instances of inadequate gating, such as in the case of weak or inconsistent respiratory signal.

When considering future directions, it is critical to understand the limitations of these experiments. First, the study was limited to use of post-mortem images to serve as motion-free controls. While histological samples would traditionally be used to confirm the presence and size of tumors, correlation of traditional 2D histological slices has resulted in confounding results. Namely, 2D pathology frequently misses microtumors, is difficult to align with 3D images, and tissue processing techniques can result in size/structural changes in the tissues that complicate volumetric validation. Thus, future studies may benefit from the addition of 3D histopathological techniques for alignment with 3D CT datasets. Implementation of the described pocket phantom was also limited by a lack of externally measured pellet sizes. To address this issue, we are pursuing the design and construction of a pocket phantom which has been machined to contain nodules of different known sizes, and specified and documented locations. With the structure and position of all pellets being defined, assessment of the phantom could be automated via scripts. These programs could then be used to quickly and systematically measure the pellets to assess the success of the gating protocols. Such a phantom could also be used to address an additional limitation of this study: comparable information from acquisitions performed on different scanners. In this way, both inter-user and inter-scanner variability might be assessed. Finally, future studies may employ more sophisticated reconstruction together with retrospective gating. Improvements to retrospective gating techniques may be able to provide image quality on par with prospectively gated images, while maintaining the advantage of shorter scan times.

Conclusions

Overall, we have demonstrated methods for respiratory gating for improving tumor detection and measurement accuracy in the lungs of free-breathing mice. Further, we have provided rigorous assessment of preclinical gating techniques, highlighting the advantages and shortcomings of multiple methods for correcting lung motion in the context of a co-clinical cancer imaging study.

Acknowledgments

Author Contributions

Conceptualization: S. J. Blocker, Y. M. Mowery, D. C. Sullivan, C. T. Badea.

Data curation: S. J. Blocker, M. D. Holbrook, C. T. Badea.

Formal analysis: S. J. Blocker.

Funding acquisition: C. T. Badea.

Investigation: S. J. Blocker, Y. M. Mowery, C. T. Badea.

Methodology: S. J. Blocker, M. D. Holbrook, D. C. Sullivan, C. T. Badea.

Project administration: Y. M. Mowery, C. T. Badea.

Resources: Y. M. Mowery, C. T. Badea.

Supervision: D. C. Sullivan, C. T. Badea.

Validation: S. J. Blocker, M. D. Holbrook.

Visualization: S. J. Blocker.

Writing – original draft: S. J. Blocker, C. T. Badea.

Writing – review & editing: S. J. Blocker, M. D. Holbrook, Y. M. Mowery, D. C. Sullivan, C. T. Badea.

References

1. Chen Z, Cheng K, Walton Z, Wang Y, Ebi H, Shimamura T, et al. A murine lung cancer co-clinical trial identifies genetic modifiers of therapeutic response. Nature. 2012; 483(7391):613–7. https://doi.org/10.1038/nature10937 PMID: 22425996

2. Nishino M, Sacher AG, Gandhi L, Chen Z, Akbay E, Fedorov A, et al. Co-clinical quantitative tumor volume imaging in ALK-rearranged NSCLC treated with crizotinib. Eur J Radiol. 2017; 88:15–20. https://doi.org/10.1016/j.ejrad.2016.12.028 PMID: 28189201

3. Ledford H. Translational research: 4 ways to fix the clinical trial. Nature. 2011; 477(7366):526–8. https://doi.org/10.1038/477526a PMID: 21956311

4. Kirsch DG, Dinulescu DM, Miller JB, Grimm J, Santiago PM, Young NP, et al. A spatially and temporally restricted mouse model of soft tissue sarcoma. Nat Med. 2007; 13(8):992–7. https://doi.org/10.1038/nm1602 PMID: 17676052

5. DuPage M, Dooley AL, Jacks T. Conditional mouse lung cancer models using adenoviral or lentiviral delivery of Cre recombinase. Nat Protoc. 2009; 4(7):1064–72. https://doi.org/10.1038/nprot.2009.95 PMID: 19561589

6. Perez BA, Ghafoori AP, Lee CL, Johnston SM, Li Y, Moroshek JG, et al. Assessing the radiation response of lung cancer with different gene mutations using genetically engineered mice. Front Oncol. 2013; 3:72. https://doi.org/10.3389/fonc.2013.00072 PMID: 23565506

7. Henschke CI, Yankelevitz DF, Yip R, Archer V, Zahlmann G, Krishnan K, et al. Tumor volume measurement error using computed tomography imaging in a phase II clinical trial in lung cancer. J Med Imaging (Bellingham). 2016; 3(3):035505.

8. Badea CT, Johnston S, Johnson B, Lin M, Hedlund LW, Johnson GA. A dual micro-CT system for small animal imaging. Proc SPIE Int Soc Opt Eng. 2008; 6913:691342. https://doi.org/10.1117/12.772303 PMID: 22049304

9. Feldkamp LA, Davis LC, Kress JW. Practical Cone-Beam Algorithm. J Opt Soc Am A. 1984; 1(6):612–9.

10. Clark D, Johnson GA, Badea CT. Denoising of 4D Cardiac Micro-CT Data Using Median-Centric Bilateral Filtration. Proc SPIE Int Soc Opt Eng. 2012;8314.

11. Sato F, Sasaki S, Kawashima N, Chino F. Late effects of whole or partial body x-irradiation on mice: life shortening. Int J Radiat Biol Relat Stud Phys Chem Med. 1981; 39(6):607–15. https://doi.org/10.1080/09553008114550731 PMID: 6972927

12. Badea CT, Drangova M, Holdsworth DW, Johnson GA. In vivo small-animal imaging using micro-CT and digital subtraction angiography. Phys Med Biol. 2008; 53(19):R319–50. https://doi.org/10.1088/0031-9155/53/19/R01 PMID: 18758005

13. Badea C, Hedlund LW, Johnson GA. Micro-CT with respiratory and cardiac gating. Med Phys. 2004; 31 (12):3324–9. https://doi.org/10.1118/1.1812604 PMID: 15651615

14. Tomasi C, Manduchi R. Bilateral filtering for gray and color images. Sixth International Conference on Computer Vision. 1998:839–46.

15. Clark D, Badea C. GPU-Based Tools for Multi-Channel X-ray CT Reconstruction2018.

16. Thirion JP. Image matching as a diffusion process: an analogy with Maxwell's demons. Med Image Anal. 1998; 2(3):243–60. https://doi.org/10.1016/s1361-8415(98)80022-4 PMID: 9873902

17. Vercauteren T, Pennec X, Perchant A, Ayache N. Diffeomorphic demons: efficient non-parametric image registration. Neuroimage. 2009; 45(1 Suppl):S61–72.

18. Johnston SM, Perez BA, Kirsch DG, Badea CT. Phase-selective image reconstruction of the lungs in small animals using Micro-CT. Proc SPIE Int Soc Opt Eng. 2010; 7622:76223G 1-G 9.

19. Holbrook M, Clark DP, Badea CT. Low-dose 4D cardiac imaging in small animals using dual source micro-CT. Phys Med Biol. 2018; 63(2):025009. https://doi.org/10.1088/1361-6560/aa9b45 PMID: 29148430

20. Van der Vorst HA. Bi-CGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of non-symmetric linear systems. SIAM Journal on Scientific and Statistical Computing. 1992; 13(2):631–44.

21. Tomasi C, Manduchi R. Bilateral filtering for gray and color images. Proc of the 1998 IEEE International Conference on Computer Vision. 1998:839–46.

22. Mai W, Badea CT, Wheeler CT, Hedlund LW, Johnson GA. Effects of breathing and cardiac motion on spatial resolution in the microscopic imaging of rodents. Magn Reson Med. 2005; 53(4):858–65. https://doi.org/10.1002/mrm.20400 PMID: 15799053

23. Kessler LG, Barnhart HX, Buckler AJ, Choudhury KR, Kondratovich MV, Toledano A, et al. The emerging science of quantitative imaging biomarkers terminology and definitions for scientific studies and regulatory submissions. Stat Methods Med Res. 2015; 24(1):9–26. https://doi.org/10.1177/0962280214537333 PMID: 24919826

24. Mulshine JL, Gierada DS, Armato SG 3rd, Avila RS, Yankelevitz DF, Kazerooni EA, et al. Role of the Quantitative Imaging Biomarker Alliance in optimizing CT for the evaluation of lung cancer screen-detected nodules. J Am Coll Radiol. 2015; 12(4):390–5. https://doi.org/10.1016/j.jacr.2014.12.003 PMID: 25842017

25. Obuchowski NA, Barnhart HX, Buckler AJ, Pennello G, Wang XF, Kalpathy-Cramer J, et al. Statistical issues in the comparison of quantitative imaging biomarker algorithms using pulmonary nodule volume as an example. Stat Methods Med Res. 2015; 24(1):107–40. https://doi.org/10.1177/0962280214537392 PMID: 24919828

26. Blocker SJ, Mowery YM, Holbrook MD, Qi Y, Kirsch DG, Johnson GA, et al. Bridging the translational gap: Implementation of multimodal small animal imaging strategies for tumor burden assessment in a co-clinical trial. PLoS One. 2019; 14(4):e0207555. https://doi.org/10.1371/journal.pone.0207555 PMID: 30958825

27. Huang J, Zhang Y, Ma J, Zeng D, Bian Z, Niu S, et al. Iterative image reconstruction for sparse-view CT using normal-dose image induced total variation prior. PLoS One. 2013; 8(11):e79709. https://doi.org/10.1371/journal.pone.0079709 PMID: 24260288