Image quality in thoracic 4D cone-beam CT: A sensitivity analysis of respiratory signal, binning method, reconstruction algorithm, and projection angular spacing

Chun-Chien Shieh

Radiation Physics Laboratory, Sydney Medical School, University of Sydney, NSW 2006, Australia and Institute of Medical Physics, School of Physics, University of Sydney, NSW 2006, Australia

John Kipritidis and Ricky T. O'Brien

Radiation Physics Laboratory, Sydney Medical School, University of Sydney, NSW 2006, Australia

Zdenka Kuncic

Institute of Medical Physics, School of Physics, University of Sydney, NSW 2006, Australia

Paul J. Keall^{a)}

Radiation Physics Laboratory, Sydney Medical School, University of Sydney, NSW 2006, Australia

(Received 16 July 2013; revised 30 January 2014; accepted for publication 4 March 2014; published 20 March 2014)

Purpose: Respiratory signal, binning method, and reconstruction algorithm are three major controllable factors affecting image quality in thoracic 4D cone-beam CT (4D-CBCT), which is widely used in image guided radiotherapy (IGRT). Previous studies have investigated each of these factors individually, but no integrated sensitivity analysis has been performed. In addition, projection angular spacing is also a key factor in reconstruction, but how it affects image quality is not obvious. An investigation of the impacts of these four factors on image quality can help determine the most effective strategy in improving 4D-CBCT for IGRT.

Methods: Fourteen 4D-CBCT patient projection datasets with various respiratory motion features were reconstructed with the following controllable factors: (i) respiratory signal (real-time position management, projection image intensity analysis, or fiducial marker tracking), (ii) binning method (phase, displacement, or equal-projection-density displacement binning), and (iii) reconstruction algorithm [Feldkamp–Davis–Kress (FDK), McKinnon–Bates (MKB), or adaptive-steepest-descent projection-onto-convex-sets (ASD-POCS)]. The image quality was quantified using signal-to-noise ratio (SNR), contrast-to-noise ratio, and edge-response width in order to assess noise/streaking and blur. The SNR values were also analyzed with respect to the maximum, mean, and root-mean-squared-error (RMSE) projection angular spacing to investigate how projection angular spacing affects image quality.

Results: The choice of respiratory signals was found to have no significant impact on image quality. Displacement-based binning was found to be less prone to motion artifacts compared to phase binning in more than half of the cases, but was shown to suffer from large interbin image quality variation and large projection angular gaps. Both MKB and ASD-POCS resulted in noticeably improved image quality almost 100% of the time relative to FDK. In addition, SNR values were found to increase with decreasing RMSE values of projection angular gaps with strong correlations ($r \approx -0.7$) regardless of the reconstruction algorithm used.

Conclusions: Based on the authors' results, displacement-based binning methods, better reconstruction algorithms, and the acquisition of even projection angular views are the most important factors to consider for improving thoracic 4D-CBCT image quality. In view of the practical issues with displacement-based binning and the fact that projection angular spacing is not currently directly controllable, development of better reconstruction algorithms represents the most effective strategy for improving image quality in thoracic 4D-CBCT for IGRT applications at the current stage.

© 2014 American Association of Physicists in Medicine. [<http://dx.doi.org/10.1118/1.4868510>]

Key words: 4D-CBCT, reconstruction, image quality

1. INTRODUCTION

In image guided radiotherapy (IGRT), the linear accelerator (linac)-mounted cone-beam computed tomography (CBCT) system plays an important role by providing 3D volumetric imaging. A CBCT scanner acquires projections over the time

scale of a few minutes because of the maximum angular speed of $6^\circ/\text{s}$ limited by the linac gantry. Consequently, images in the thoracic region are subject to appreciable motion artifacts due to breathing. Respiratory-correlated 4D-CBCT is therefore required for accurate monitoring of the tumor trajectory.¹ 4D-CBCT is achieved by sorting projections into bins

(respiratory binning) corresponding to different motion phases estimated by a specific respiratory signal, and reconstructing each bin, respectively. Thus, respiratory signal, respiratory binning, and reconstruction algorithm are the three controllable factors that affect the resultant image quality. In addition, since the reconstruction is based on information from various projection angular views, projection angular spacing is also an important factor passively affected by different respiratory signals and binning methods.

Respiratory signals can be obtained by tracking the implanted radio-opaque fiducial markers in the projection images, which provides the most reliable correlation with internal motion.^{2,3} A robust template-based marker segmentation algorithm has been proposed for such tracking, and was shown to succeed in more than 99.9% of the CBCT cases as long as at least one marker is visible in all projection images.⁴ However, this is generally not clinically used as implanted markers are not common to all patients, and marker migration is also a potential problem.⁵ External surrogates are considerably more clinically viable. The Varian real-time position management (RPM) system (Varian Medical Systems, Palo Alto, CA), which tracks patient abdominal motion, is widely used as an external respiratory signal source. Nevertheless, there are concerns regarding potential mismatch between external and internal motions.^{6,7} Respiratory signals can also be extracted directly from the projections via the intensity analysis (IA) method,^{8,9} the Amsterdam shroud (AS) method,^{10,11} the Fourier-transform based phase analysis method,¹² or the local principal component analysis method.¹³ It has been shown that the IA method is slightly less accurate but more stable than the others, whereas the AS method provides the best accuracy when distinct moving objects are present in the projection views.¹³ Currently, the AS method is clinically used by the Elekta Synergy 4D-CBCT system (Elekta Oncology Systems Ltd, Crawley, UK). In practice, internal surrogates, external surrogates, and projection based methods may result in different outcomes. This is the first study that compares the impact of these competing methods on image quality.

For respiratory binning, the two common approaches are phase binning (PB) (Refs. 14 and 15) and displacement binning (DB).^{16,17} The former sorts projections according to the phase of the respiratory signal in each breathing cycle while the latter groups projections according to the displacement of the respiratory signal. Several studies have shown that displacement binning results in fewer motion artifacts in 4D-CT, and recovers the sizes and shapes of moving objects more accurately,^{17–19} but no similar study for 4D-CBCT exists to date. The disadvantage of displacement binning is that it often leads to missing CT slices either due to baseline shifts or the intrinsic nonlinearity of the breathing motion pattern. In CBCT, this is likely to cause uneven distribution of projections angular views in each bin, strongly affecting the outcomes of the reconstruction. In this study, the effect of displacement binning on 4D-CBCT reconstruction is investigated for the first time.

At the moment, CBCT images are reconstructed using the Feldkamp–Davis–Kress (FDK) algorithm²⁰ in practice,

which is an approximated filtered backprojection. Despite its computational efficiency, it is prone to streaking and noise in 4D-CBCT, as projection angular sampling is usually low in most respiratory bins. A common approach for streaking and noise reduction is the McKinnon–Bates (MKB) algorithm,^{21,22} which involves the smooth yet temporally blurred 3D standard image reconstructed without respiratory binning as an initial image. A correction image estimated from the difference between the original projection data and the forward projections of the standard image is then added to correct for motion artifacts. However, the improvement in image quality may be limited.²³ Furthermore, it may suffer from residual motion artifacts due to the use of the standard image. The emergence of compressed sensing theory²⁴ provided 4D-CBCT reconstruction with a new possibility. To reduce noise and streaking, compressed sensing based algorithms sparsify an alternative representation of the image, usually iteratively. In particular, total-variation (TV) minimization (sparsification of the image gradient) has been shown to be effective in noise and streaking suppression. The adaptive-steepest-descent projection-onto-convex-sets (ASD-POCS) (Ref. 25) algorithm is one of the most popular TV-minimization based methods. It employs an algebraic reconstruction technique (ART) to maintain data fidelity while reducing the noise with a steepest-descent TV-minimization step in every iteration. In spite of the much lower level of noise and streaking, ASD-POCS appears to cause loss of contrast on fine anatomical structures.²³

Another reconstruction factor crucial to image quality is the angular spacing between projections. Although one has no direct control over the angular spacing, it is affected by the use of different respiratory signals and respiratory binning methods. It has been shown that projection clustering leads to unnecessary patient dose that does not improve image quality.²⁶ Large and irregular angular spacing was also shown to degrade images.²⁷ This is the first study to investigate how angular spacing affects image quality, and how it is related to the choice of respiratory signal, binning method, and reconstruction algorithm.

Although various “intrafactor” comparisons have been conducted, no “interfactor” comparison exists to date. The purpose of this study is to: (i) investigate the sensitivity of image quality to the three controllable factors; (ii) investigate how projection angular spacing affects image quality, and how it is related with the three controllable factors. Overall, this study aims to find the most important factors to image quality in thoracic 4D-CBCT.

2. MATERIALS AND METHODS

2.A. Patient projection data

The thoracic CBCT half-fan projection datasets of seven patients acquired with audiovisual biofeedback from the Varian On-Board Imager (Varian Medical Systems, Palo Alto, CA) were used. The patients were from an NCI-sponsored lung cancer trial because at least one gold marker was implanted in every patient.²⁸ The implanted markers are

essential to this study for two reasons: (i) using marker tracking as a respiratory signal source (see Sec. 2.B.1.c); (ii) using the marker as a surrogate for measuring image quality (see Sec. 2.D.2). There were two scans for each patient with treatments in between, giving a total of 14 projection datasets. The time interval between the two scans for each patient ranged from two days to approximately two months. It should be noted that the treatment and time interval between the two scans are not pertinent to this study, since none of the analysis was directly related to tumors. Instead, including two scans for each patient allows investigation of both “inpatient” and “interpatient” image quality differences, which helps distinguish between patient specific and random variations.

In order to compare the impact of different respiratory signal sources, at least one gold marker must be visible in all projection frames, so that marker tracking can be used as a source of respiratory signal for the particular scan. Out of all 14 scans (7 patients), only scans 1–6 (patients 1–3) fulfill this prerequisite, and were thus used to investigate the impact of respiratory signal sources. The impact of respiratory binning, reconstruction algorithm, and projection angular spacing were studied using all 14 scans.

The imaging x-ray configuration was 125 kVp, 20 mA, and 20 ms per projection. The respiratory signals acquired from RPM, IA, and marker tracking (outlined in Sec. 2.B.1) for scans 1–6 are displayed in Fig. 1. Respiratory signals for

FIG. 1. The (a) RPM, (b) intensity analysis, and (c) marker signals for scans 1–6 from patient 1, 2, and 3 (two scans per patient). The respiratory signals for scans 7–14 show similar variability, and are not displayed. All data were plotted with respect to projection number on the x -axis, with peaks representing end inhalation. The RPM and marker signals were plotted in millimeters. The intensity analysis signals do not represent absolute displacement, and were plotted in arbitrary units (AU). For each scan, 1 AU represents the standard deviation of the signal. The mean values of all the data were shifted to zero. The dotted lines in (a) and (c) represent the upper and lower boundaries in displacement binning defined in Sec. 2.B.2.b.

TABLE I. Mean ranges of motion, baseline shifts, respiratory periods, and phase shifts presented in scans 1–6. Scans 7–14 show similar variability in respiratory motion features, and are not included in this table. The baseline shifts were assessed by the RMSE of baseline positions. Respiratory periods were defined to be the mean time interval between consecutive troughs, the RMSE values of which represent phase shifts. The IA signal does not represent absolute displacement, and was not used to calculate mean amplitudes and baseline shifts.

Patient Scan		1		2		3	
		1	2	3	4	5	6
Mean motion range (mm)	RPM	7.8	5.1	5.0	5.5	12.2	16.8
	Marker	4.5	3.4	3.1	2.7	7.0	7.1
Baseline shift (mm)	RPM	0.4	0.3	0.5	0.5	0.5	0.9
	Marker	0.2	0.3	0.4	0.3	0.5	0.4
Mean respiratory periods (s)	RPM	4.2 ± 1.3	4.5 ± 1.2	3.9 ± 1.1	3.4 ± 1.1	4.1 ± 1.5	4.1 ± 1.6
	IA	4.0 ± 1.0	4.2 ± 1.1	3.8 ± 1.0	3.2 ± 1.0	3.9 ± 1.3	4.2 ± 1.6
	Marker	4.2 ± 1.2	4.1 ± 0.7	3.7 ± 1.0	3.5 ± 1.1	4.3 ± 1.5	3.7 ± 1.0

scans 7–14 show similar variability, and are not shown here. A range of different motion amplitudes, baseline shifts, respiratory periods, and phase shifts are observed not only between different scans, but also between different respiratory signals (see Table I), showing that some disparity in motion estimation exists for different respiratory signal sources. All 12 scans each consisted of about 2400 projections spanning an angular range of 360° . With a frame rate of about 4.9 fps, the acquisition time was 7–9 min for each scan, during which approximately 100 respiratory cycles were observed. The detector was laterally displaced by 14.8 cm for half-fan acquisition. The dimension of projection images is 1024×768 with a pixel spacing of 0.388 mm in both lateral and longitudinal directions.

2.B. Controllable factors

For each scan, images were reconstructed multiple times varying the following three controllable factors: respiratory signal, binning method, and reconstruction algorithm. The details of the various options for each factor are described in this section.

2.B.1. Respiratory signal

RPM signal, projection based intensity analysis method, and fiducial marker tracking were used to estimate the internal organ motion due to breathing, and are outlined as follows.

2.B.1.a. Real-time position management (RPM). The Varian RPM system (Varian Medical Systems, Palo Alto, CA) tracks the position of a reflective marker positioned on a patient's abdomen. The displacement of the abdominal motion is recorded at a rate of approximately 30 Hz. Since this is higher than the frame rate of the CBCT imager, the data were down-sampled to match the imaging frequency.

2.B.1.b. Intensity analysis. The IA method^{8,9} was used to extract respiratory signal directly from the projection images because it has the most stable performance.¹³ Respiratory motion is the most evident near the diaphragm and the lower lobes, which can be observed in the lower-left region of

the projection images in half-scan acquisition. Consequently, the raw data were taken as the sum of the attenuation values in the lower-left region of each projection image. A high-pass filter was then applied to separate the variation due to respiratory motion (high frequency) and the gantry rotation (low frequency).

It should be noted that IA signal does not represent absolute displacement because the gantry rotation component cannot be removed completely. Therefore, IA signal was only used for phase binning in this study.

2.B.1.c. Fiducial marker tracking. Scans 1–6 each contain at least one fiducial marker visible in all projection frames. For these scans, the marker around the central area of the thoracic cavity that is the most visually distinct in the projection views was segmented using a robust template-based segmentation method.⁴ 3D trajectories of the markers were then calculated, and the superior–inferior (SI) components were used as the respiratory signals. This method was shown to obtain successful segmentations in more than 99.9% of the CBCT segmentation cases with multiple fiducial markers. However, in very few projection frames, the marker could be unreliably segmented when covered by dense structure.

2.B.2. Respiratory binning method

For each respiratory signal described in Sec. 2.B.1, projection images were sorted into ten respiratory bins using phase, displacement, or equal-projection-density displacement binning methods. Details of each are outlined below.

2.B.2.a. Phase binning. For phase binning, respiratory cycles were first identified by locating subsequent troughs (end exhalations) in the respiratory signal. For each cycle, the phase is piecewise-linear, with end exhalation fixed at 0 and 2π , and end inhalation fixed at π .

2.B.2.b. Displacement binning. The displacement of the signal was divided into six equally spaced sections. Except for the highest (end inhalation) and lowest (end exhalation) sections, projection views in each section were classified to be in midinhalation or midexhalation depending on whether the gradient was positive or negative.

It is worth mentioning that using the global maximum and minimum of the respiratory signal as the upper and lower boundaries for division may cause large angular gaps when there are considerable baseline shifts. Therefore, inspired by the respiratory extrema selection method for displacement-based 4D-CT proposed by Fitzpatrick *et al.*,¹⁷ we defined the upper and lower boundaries to be the median values of all the peak and trough positions. Any displacements beyond/below the upper/lower boundary were designated to the end inhalation/end exhalation bins. The upper and lower boundaries are represented by black dotted lines in Figs. 1(a) and 1(c).

2.B.2.c. Equal-projection-density displacement binning (EPDB). Displacement binning often results in uneven allocation of projection images between respiratory bins, because although the displacement respiratory bins are spaced evenly in displacement, the corresponding time intervals can be very different. To address uneven allocation of projection images, the alternative “equal-projection-density displacement binning” works as follows.

1. The highest (end inhalation) and lowest (end exhalation) displacement sections were determined such that they were each allocated one-tenth of the total number of projection images.
2. The rest of the projection images were categorized into midinhalation or midexhalation depending on the gradient of the respiratory signal.
3. The midinhalation projection images were divided into four unequally spaced respiratory bins such that the number of projection images in each bin was similar.
4. The midexhalation projection images were sorted into four respiratory bins in the same way, resulting in a total of $2 + 4 + 4 = 10$ respiratory bins.

It is worth mentioning that although EPDB assures a more consistent number of projection images in every respiratory bin, the projection angles in each bin may still be clustered. Thus, EPDB does not guarantee even distribution of projection angles.

2.B.3. Reconstruction algorithm

After respiratory binning, each bin was reconstructed using FDK, MKB, or ASD-POCS. These three algorithms have been comprehensively described by Bergner *et al.*²³ In summarizing their descriptions below, we denote the reconstructed image as $\tilde{f}$, the projection data as $\tilde{p}$, and the forward projection operator as R .

2.B.3.a. FDK. The FDK algorithm²⁰ is essentially an approximated filtered backprojection of 2D projection data:

$$\tilde{f}_{\text{FDK}} = R_{\text{FDK}}^{-1} \tilde{p}. \quad (1)$$

2.B.3.b. MKB. The MKB algorithm^{21,22} utilizes the standard image $\tilde{f}_{\text{Std}}$ (reconstructed without respiratory binning) as the initial image. The difference between the original projection data and the forward projection of $\tilde{f}_{\text{Std}}$ is then

calculated. Backprojection of this “difference projection” is added to $\tilde{f}_{\text{Std}}$ to correct for motion artifacts. This can be written as

$$\tilde{f}_{\text{MKB}} = \tilde{f}_{\text{Std}} + R_{\text{FDK}}^{-1} (\tilde{p} - R \tilde{f}_{\text{Std}}). \quad (2)$$

2.B.3.c. ASD-POCS. ASD-POCS (Ref. 25) is a compressed sensing based iterative algorithm. It consists of two components: the ASD component and the pPOCS component. The ASD component minimizes the total-variation (TV) of the image by a gradient descent step, but in general decreases data fidelity. The POCS component, which utilizes an algebraic-reconstruction-technique (ART) step with positivity constraint, is therefore employed to maintain data fidelity. Both components exploit adaptive step sizes to prevent overshooting. Alternations between ASD and POCS solve for a solution that agrees with the projection data within an error tolerance ϵ with minimized TV. The optimization problem can be expressed as

$$\tilde{f}_{\text{ASD-POCS}} = \underset{\tilde{f}}{\operatorname{argmin}} \|\tilde{f}\|_{\text{TV}}, \quad \text{subject to } \|\tilde{p} - R\tilde{f}\|^2 < \epsilon^2. \quad (3)$$

The convergence criterion is such that c_α , defined as the cosine of the angle between the step changes due to the ASD and the POCS components $\Delta \tilde{f}_{\text{ASD}}$ and $\Delta \tilde{f}_{\text{POCS}}$:

$$c_\alpha = \frac{\Delta \tilde{f}_{\text{ASD}} \cdot \Delta \tilde{f}_{\text{POCS}}}{\|\Delta \tilde{f}_{\text{ASD}}\| \|\Delta \tilde{f}_{\text{POCS}}\|}, \quad (4)$$

is sufficiently close to negative one. This is when $\Delta \tilde{f}_{\text{ASD}}$ and $\Delta \tilde{f}_{\text{POCS}}$ point in opposite directions in the solution space, and further iterations merely result in the oscillation of $\tilde{f}$ around the desired solution.

2.C. Reconstruction parameters

Scans 1–6 were each reconstructed multiple times varying all three controllable factors ([RPM, IA, marker tracking] $\times$ [PB, DB, EPDB] $\times$ [FDK, MKB, ASD-POCS]), while scans 7–14 were reconstructed using only the RPM signals ([RPM] $\times$ [PB, DB, EPDB] $\times$ [FDK, MKB, ASD-POCS]). For each 4D reconstruction, the projections were sorted into ten respiratory bins. The reconstructed image space contained 450^2 voxels in 100 axial slices. Each voxel was 1 mm^2 in size, and each slice was spaced 2 mm apart. Before reconstruction, every projection image was downsampled from a resolution of 0.388 mm (both lateral and vertical) to 1.552 mm (lateral) and 3.104 (vertical) by cubic interpolation, which matches the reconstruction resolution at isocenter (1mm in both coronal and sagittal direction, and 2 mm in the axial direction) when multiplied by the ratio SAD/SDD $\approx 2/3$. Although this may reduce image resolution at regions far from the isocenter, the regions-of-interest are generally not noticeably affected as they are reasonably close to the isocenter. Downsampling of projection images not only speeds up the reconstruction, but also reduces Poisson noise since multiple

projection pixels are rebinned into one.²⁹ All reconstructed images were cropped to the CBCT field-of-view. The images were reconstructed in attenuation values (mm^{-1}) with a positivity constraint.

For ASD-POCS, the ART step was realized as SART instead. The backprojection in SART was approximated by FDK for better computational efficiency without noticeable loss of accuracy.³⁰ The relaxation reduction factor β_{red} was set to 0.7, the steepest-descent reduction factor α_{red} was set to 0.6, and the threshold r_{max} for changing the step size of the ASD component was set to 0.2. A detailed description of all the parameters is given in Ref. 25. In practice, the convergence of $c_\alpha = -1$ is often difficult to achieve. Therefore, the iteration was ceased once c_α changed by less than 10^{-3} .²³

FDK reconstructions were carried out using Reconstruction Toolkit developed by Simon Rit.³¹ The forward projection and FDK backprojection modules in Reconstruction Toolkit were adapted into our in-house MATLAB codes for MKB and ASD-POCS reconstructions.

2.D. Image quality metrics

Regarding the usability of 4D-CBCT images in IGRT, four different types of image degradation are of major concern: (i) motion blur, (ii) interbin image quality variation, (iii) noise and streaking, and (iv) blur caused by oversmoothing from the reconstruction algorithm. The former two are related to projection allocation in respiratory bins, and are governed by the source of respiratory signal and binning method, while the latter two are governed by the choice of reconstruction algorithm. For clinical data, it is difficult to assess image degradation directly because of the lack of ground truth. Consequently, image quality metrics are often used as surrogates for measuring such degradation.

In this study, image quality was assessed using signal-to-noise ratio (SNR), contrast-to-noise ratio (CNR), and edge-response width (ERW). The mean and root-mean-squared-error (RMSE) values of each metric over all respiratory bins were compared between each controllable factor. The calculation of each metric was specifically defined for this study in order to assess the four major types of degradation and to allow comparisons between the three controllable factors. Denoting pixel intensity as I , the detailed definition of each metric is outlined in Secs. 2.D.1–2.D.3. The relations between the three controllable factors, the three image quality metrics, and the four major types of image degradation are summarized in Table II.

2.D.1. Signal-to-noise ratio

SNR measures the smoothness of the image, taking both the effects of noise and streaking into account. A lower SNR value indicates higher level of noise and streaking, which can potentially degrade the visibility of important anatomical details. In this study, SNR was defined over a smooth soft tissue region R between and well away from the two lungs (see

TABLE II. The four major types of 4D-CBCT image degradation that are of clinical interest. These degradations are either related to the use of different respiratory signals and binning methods, or different reconstruction algorithms. A tick indicates that mean or RMSE value of the particular image quality metric is an effective surrogate for measuring the corresponding type of image degradation.

Image degradation	SNR	CNR	ERW
Respiratory signal and respiratory binning method			
Motion blur		✓	✓
Interbin image quality variation	✓	✓	
Reconstruction algorithm			
Noise and streaking	✓	✓	
Blur caused by oversmoothing		✓	✓

Fig. 2). The volume of the region R was approximately 1cm^3 , with slight variation between patients. SNR was then calculated by

$$\text{SNR} = \frac{\text{mean}(I_R)}{\text{RMSE}(I_R)}. \quad (5)$$

SNR values are the most pronouncedly affected in the following ways.

1. A noise/streaking reduction reconstruction algorithm increases mean SNR values.
2. A different respiratory signal and binning method result in a different distribution of projection angular views in each respiratory bin, thereby affecting the RMSE values of SNR (interbin image quality variation).

It should be noted that although motion blur and oversmoothing may also affect SNR values, the smooth soft tissue region R was selected such that it was well away from the lung boundaries, and contained no contrast object. As a result, the effects of motion blur and oversmoothing on SNR values are minimal in this study.

FIG. 2. The homogeneous region R over which SNR was defined, and the fiducial marker used for CNR calculation.

2.D.2. Contrast-to-noise ratio

CNR measures the visibility of small high contrast objects. A lower CNR value indicates worse visibility of anatomical details. The intensities of the gold marker and surrounding tissues (region R) as shown in Fig. 2 were used to define CNR, given by

$$\text{CNR} = \frac{I_{\text{marker}} - \text{mean}(I_R)}{\text{RMSE}(I_R)}. \quad (6)$$

The denominator in Eq. (6) accounts for noise and streaking, while the numerator accounts for blur caused by oversmoothing by reconstruction algorithm or motion artifacts. Figure 3 illustrates the reduction in marker CNR values due to motion blur. In the reconstruction of a motion blurred image, the marker is backprojected to a wide neighborhood along its motion trajectory, causing its pixel intensity to be smeared over a larger region than in the respiratory-correlated case. This reduces the pixel intensity and results in a lower CNR value.

CNR values are the most pronouncedly affected in the following ways.

1. A reconstruction algorithm may either increase mean CNR values due to the reduction in noise/streaking, or decrease mean CNR values due to oversmoothing the object.
2. A different respiratory signal and binning method result in a different distribution of projection angular views in each respiratory bin. This affects the noise and streaking level in each bin, and therefore the RMSE values of CNR (interbin image quality variation).
3. An inaccurate respiratory signal or inappropriate binning method results in more motion blur, thereby decreasing mean CNR values.

FIG. 3. CBCT axial image slices illustrating the effect of motion blur on marker CNR values. The respiratory-correlated image (left) was reconstructed with marker signal, displacement binning, and FDK reconstruction. The motion blurred image (right) was reconstructed using FDK with all the projection images. In each case, the axial slice with the best marker visibility was selected. It can be seen clearly that the contrast of the marker in the motion blurred image is much worse than that in the respiratory-correlated image. The window level was set to $[0, 0.04] \text{ mm}^{-1}$ (attenuation coefficient) in order to show the difference between marker intensities.

2.D.3. Edge-response width

ERW measures the boundary sharpness between different regions. A higher ERW value represents worse boundary sharpness. Figure 4 illustrates how ERW was defined. The linear intensity profile across the lower middle boundary of the left lung was first acquired. The mean lung and tissue intensities represent 0% and 100% of the intensity change at the boundary. The ERW was defined as the spatial width between the 25% and 75% intensity change.

Similar to CNR, ERW accounts for both oversmoothing by the reconstruction algorithm and motion blur. Figure 5 illustrates how the lung boundary is smeared along its motion trajectory in the reconstruction of a motion blurred image, leading to a larger ERW value than in the respiratory-correlated case.

ERW values are the most pronouncedly affected in the following ways.

1. A reconstruction algorithm may increase mean ERW values due to oversmoothing the boundary.
2. An inaccurate respiratory signal or inappropriate binning method results in more motion blur, thereby increasing mean ERW values.

ERW is in general not as sensitive to noise and streaking as CNR for the following reasons. First, intensity changes across the lung boundary are usually much steeper and more distinct than random variations in pixel intensity due to noise. Second, the linear intensity profiles were acquired at locations where no apparent streak was involved. As a result, noise and streaking have minimal effects on ERW values in this study.

2.E. Projection angular spacing

Respiratory binning in 4D-CBCT often leads to inappropriate projection angular spacing, which results in noise and

FIG. 4. CBCT axial image slice and linear intensity profile showing the definition of ERW used in this study: the 25%–75% width of the linear intensity profile across the lung boundary.

FIG. 5. CBCT axial image slices and linear intensity profiles illustrating the effect of motion artifacts on ERW values of the lung boundary. The respiratory-correlated image (left) was reconstructed with marker signal, displacement binning, and FDK reconstruction. The motion blurred image (right) was reconstructed using FDK with all the projection images. An axial slice with clear lung motion was selected. It can be seen clearly that the lung boundary in the motion blurred image is less sharp than that in the respiratory-correlated image, resulting in a larger ERW value. The window level was set to $[0, 0.02] \text{ mm}^{-1}$ (attenuation coefficient).

streaking, thereby low SNR values. To investigate how image quality is affected by projection angular spacing $\Delta\theta$, the SNR value of every reconstructed image in each scan was plotted with respect to three different angular spacing measurements: $\max(\Delta\theta)$, $\text{mean}(\Delta\theta)$, and $\text{RMSE}(\Delta\theta)$. The maximum angular gap represents the largest region uncovered by projection data. The mean angular gap is essentially the same as the reciprocal of the number of projection images. The RMSE of angular gaps describes the evenness of projection angular distribution. The correlation coefficient r between SNR and each of the measurement, expected to be negative as larger angular spacing results in worse image quality,²⁴ was calculated to determine the most sensitive angular spacing factor. In addition, to investigate the effects of respiratory signal and binning method on projection angular spacing, the distribution of all three angular spacing factors was also analyzed with respect to different respiratory signals and binning methods.

3. RESULTS

3.A. Qualitative comparison

Figure 6 displays the axial views of end inhalation images reconstructed using different respiratory signals, bin-

ning methods, and reconstruction algorithms (patient 1, scan 1). Images from the other respiratory phases, scans and patients show similar trends. It is evident that reconstruction algorithms cause distinguishable image qualities. FDK reconstruction generated the most noise and streaking, followed by MKB, then ASD-POCS. The MKB images are far more visually acceptable compared to FDK, and show no observable blurring effect. Although a considerable amount of noise and streak artifacts remain, the MKB images show better overall contrast. The reduction in noise and streaking from ASD-POCS reconstruction is the most significant, but at the cost of losing contrast in the bony anatomy due to oversmoothing. However, ASD-POCS showed the best marker visibility, followed by MKB and FDK. Motion details such as lung boundaries, small vessels, and fiducial markers are well preserved in ASD-POCS images.

In terms of motion artifacts, there are no observable qualitative differences between the uses of different respiratory signals, despite the observable differences in motion features and patterns they showed (see Table I and Fig. 1). Similarly, although respiratory motions are treated very differently in different binning methods, it is difficult to conclude any apparent differences in motion details from qualitative investigations of the resultant images.

FIG. 6. Axial views of end inhalation images (scan 1, patient 1) reconstructed using different respiratory signals, respiratory binning methods, and reconstruction algorithms. The marker was located near the inner side of the right lung. The window level was set to $[0, 0.02] \text{ mm}^{-1}$ (attenuation coefficient). (PB: phase binning; DB: displacement binning; EPDB: equal-projection-density displacement binning; IA: intensity analysis).

3.B. Image quality metrics vs three controllable factors

In Fig. 7, the mean SNR, CNR, and ERW values for patient scans 1–6 are plotted with respect to different respiratory signals, binning methods, and reconstruction algorithms for comparison between their impacts on image quality. The error bars represent the RMSE values over the ten bins of each 4D reconstruction. Similar trends between binning methods and reconstruction algorithms were observed for scans 7–14 (not shown here). For the interpretation of results in Fig. 7, the reader is referred to Table II, where possible effects of controllable factors on the three image quality metrics are outlined. For different respiratory signals and binning methods, the mean CNR and ERW values (motion blur), as well as the RMSE values of SNR and CNR (interbin image quality variation) are of interest. For different reconstruction algorithms, the mean SNR values (noise and streaking), mean CNR values (noise, streaking, and oversmoothing), and mean ERW values (oversmoothing) are of interest.

3.B.1. Respiratory signal

Respiratory signals have negligible effect on either the mean CNR and ERW values, or the RMSE values of SNR and CNR. Except for a significantly lower mean CNR value in scan 1 using IA signal and ASD-POCS reconstruction, which was not repeatedly observed in any other scans, all differences are negligible. Moreover, variations in metric values between different respiratory signals show neither consistent interpatient nor inpatient trend. The lack of interpatient trend suggests that none of the three sources of respiratory signals performs consistently better than the other two. The lack of inpatient trend suggests that the performances of different respiratory signals are not noticeably dependent on patient specific factors (e.g., breathing habit). Therefore, differences in the image quality metrics corresponding to different respiratory signals can most likely be attributed to random variations. In other words, based on the results of this three-patient study (patients 1–3), respiratory signals were found to have no observable impact on (i) motion artifacts (similar CNR and ERW values), and (ii) interbin image quality variations (similar RMSE values of SNR and CNR). However, these conclusions need to be further verified in a larger patient cohort.

3.B.2. Respiratory binning method

Figure 7(b) shows that displacement binning and equal-projection-density displacement binning (displacement-based binning) produced notably higher CNR values than phase binning, except in patient 1, scan 2 where all binning methods had similar performances. On the other hand, much smaller and less consistent differences in mean SNR values were observed between phase binning and displacement-based binning, indicating that binning methods have no significant impacts on the overall noise and streaking level as expected. Since displacement-based binning resulted in similar noise and streaking level relative to phase binning, the superior CNR values can be solely attributed to less motion blur (as discussed in Sec. 2.D.2). The higher CNR values were not obvious in Fig. 6 since the marker intensity in a respiratory-correlated image is generally much higher than the window value for visualization. Nevertheless, as the marker CNR values act as surrogates for measuring motion blur, this result indicates that displacement-based binning is less prone to motion artifacts than phase binning regarding small anatomical details. On the other hand, all three respiratory binning methods resulted in very similar ERW values [see Fig. 7(c)], showing no observable differences in motion artifacts regarding boundary sharpness under the resolution used in this study ($1 \times 1 \times 2 \text{ mm}$).

Figures 7(a) and 7(b) also show that displacement-based binning methods usually resulted in higher RMSE values of both SNR and CNR than phase binning, therefore larger interbin image quality variation. It is also noteworthy that equal-projection-density displacement binning did not lead to smaller interbin image quality variations, even though it aims for even allocation of projection images in all respiratory bins.

FIG. 7. Mean (a) SNR, (b) CNR, and (c) ERW values over ten respiratory bins for scans 1–6 plotted against different respiratory signals and binning methods using different reconstruction algorithms. The error bars represent the RMSE of the metric values over ten respiratory bins. The black dashed lines separate results obtained from different binning methods. Note that higher SNR, CNR values, and lower ERW values represent better image quality. Therefore, (c) ERW values were plotted on an inverted axis such that an upward trend represents better image quality on all of the graphs. (PB: phase binning; DB: displacement binning; EPDB: equal-projection-density displacement binning; IA: intensity analysis.)

Both observations are related to projection angular spacing, and will be discussed further in Sec. 3.C.2.

To further verify the observations from Fig. 7, the relative performance of displacement-based binning and phase binning in terms of SNR, CNR, and ERW was compared using Student's *t*-tests on every 4D reconstruction case from all 14 scans. The reader is referred to Appendix D for more details on the Student's *t*-tests. Figure 8(a) displays the percentages of 4D reconstruction cases in which displacement-based binning performed better/worse or similar relative to phase binning. Displacement-based binning and phase binning produced similar SNR values in the majority of cases studied. At other times, displacement-based binning did not lead to consistently higher or lower SNR values. In contrast, displacement-based binning resulted in higher CNR values than phase binning in more than half of the cases, concurring with the observation that displacement-based binning resulted in notably improved motion sharpness. Regarding boundary sharpness, displacement-based binning produced lower ERW values (better sharpness) more often than higher ERW values

relative to phase binning. Nevertheless, over 75% of the time there was no significant difference in ERW values between displacement-based binning and phase binning.

3.B.3. Reconstruction algorithm

Figure 7(a) shows that different reconstruction algorithms lead to distinctly different SNR values. ASD-POCS performed significantly better than the other two, and MKB performed slightly but consistently better than FDK. The same trend occurred in CNR [see Fig. 7(b)]. Although there were variations in CNR values from different respiratory binning methods, ASD-POCS always resulted in the highest CNR values, followed by MKB, then FDK. From the reconstructed images in Fig. 6, it is clear that the improvements in CNR values from MKB and ASD-POCS result from the suppression of noise and streaking without oversmoothing the fiducial markers or other motion details. It is noteworthy that although qualitative analysis (cf. Fig. 6) shows

FIG. 8. (a) Stacked bar charts showing percentages of 4D reconstruction cases in which DB and EPDB performed better, similar, or worse compared to PB in terms SNR, CNR, and ERW. (b) Stacked bar charts showing percentages of 4D reconstruction cases in which MKB and ASD-POCS performed better, similar, or worse compared to FDK in terms SNR, CNR, and ERW. It should be noted that better image quality corresponds to higher SNR, higher CNR, and lower ERW values. The comparisons were conducted using Student's *t*-tests on every 4D reconstruction from all 14 scans. The reader is referred to Appendix D for more details on the Student's *t*-tests.

slight blurring in the bony anatomy from ASD-POCS, most motion details remained reasonably sharp (see Appendix B for example images). On the other hand, slightly higher ERW values (worse boundary sharpness) were observed for ASD-POCS than MKB and FDK [cf. Fig. 7(c)], indicating that ASD-POCS may cause tiny but observable edge blur.

The relative performance of MKB and ASD-POCS was compared to FDK in terms of SNR, CNR, and ERW using Student's *t*-tests on every 4D reconstruction cases from all 14 scans (see Appendix D for details). Figure 8(b) displays the percentages of 4D reconstruction cases in which MKB or ASD-POCS performed better/worse or similar compared to FDK. MKB and ASD-POCS led to statistically significant improvements in both SNR and CNR values for 98% and 100% of the cases, respectively. Regarding boundary sharpness, MKB produced similar ERW values to FDK in the majority of cases studied. In contrast, ASD-POCS produced higher ERW values than FDK in approximately 40% of the cases, indicating that edge blur caused by ASD-POCS is small but observable.

3.C. SNR vs projection angular spacing

3.C.1. Comparison between the projection angular spacing and image quality

To investigate the effect of projection angular spacing on image quality, SNR values of FDK images in all ten respiratory bins reconstructed using different respiratory signals and binning methods in scans 1–6 were plotted with respect to $\max(\Delta\theta)$, $\text{mean}(\Delta\theta)$, and $\text{RMSE}(\Delta\theta)$ in Fig. 9, with the correlation coefficient r calculated for each case. Results from scans 7–14 show similar trends, and are not displayed. Figure 9(a) shows that SNR values were reasonably correlated with $\max(\Delta\theta)$, but exceptionally low correlations existed for some cases (e.g., scan 5, where $r = -0.00$). Figure 9(b) shows that there is a consistent but small negative correlation between $\text{mean}(\Delta\theta)$ and SNR values. The best correlation was observed between SNR and $\text{RMSE}(\Delta\theta)$ as shown in Fig. 9(c), implying that image quality is the most sensitive to $\text{RMSE}(\Delta\theta)$.

The mean correlation coefficients over all 14 scans are summarized in Table III. The variability from scan to scan was

FIG. 9. SNR values of images reconstructed using FDK plotted with respect to (a) $\max(\Delta\theta)$, (b) $\text{mean}(\Delta\theta)$, and (c) $\text{RMSE}(\Delta\theta)$ for scans 1–6. Results for scans 7–14 show similar trends and are not displayed. Different respiratory signals and binning methods are represented by different shapes (RPM: circle; intensity analysis: square; marker: cross) and gray levels (phase: black; displacement: dark gray; equal-projection-density displacement: light gray), respectively. The correlation coefficient r for each case is also displayed.

estimated by the standard deviation. In general, SNR values were strongly and consistently correlated with $\text{RMSE}(\Delta\theta)$ ($|r| \geq 0.69$), while the correlations with $\text{mean}(\Delta\theta)$ were much smaller. SNR values were reasonably correlated with $\max(\Delta\theta)$ ($|r| \geq 0.4$), but the correlation coefficients varied

much more from scan to scan than the other two projection angular spacing factors.

Table III also shows that reconstruction algorithms may mitigate the impact of projection angular spacing on image quality. For example, ASD-POCS reduced the mean correlation coefficient between SNR and $\text{RMSE}(\Delta\theta)$ from -0.78 to -0.69 . This confirms expectations that with a better reconstruction algorithm, lower or less ideal projection angular sampling may be tolerable. Nevertheless, a correlation coefficient of -0.69 still implies a significant impact of uneven projection angular distribution on image quality. This suggests that acquisition of even projection angular views is crucial to reducing noise and streaking even when a noise/streaking-reducing reconstruction algorithm is employed.

3.C.2. Interbin image quality variation

It can be seen in Fig. 9(c) that displacement binning (dark gray) and equal-projection-density displacement

TABLE III. Correlation coefficients r between SNR values and the projection angular spacing factors, $\max(\Delta\theta)$, $\text{mean}(\Delta\theta)$, and $\text{RMSE}(\Delta\theta)$, calculated for different reconstruction algorithms. The recorded values are the average correlation coefficients over all 14 scans, and the variability from scan to scan was estimated by the standard deviation.

Reconstruction algorithm	SNR vs $\max(\Delta\theta)$	SNR vs $\text{mean}(\Delta\theta)$	SNR vs $\text{RMSE}(\Delta\theta)$
FDK	-0.46 ± 0.27	-0.32 ± 0.10	-0.78 ± 0.06
MKB	-0.41 ± 0.28	-0.31 ± 0.10	-0.78 ± 0.07
ASD-POCS	-0.40 ± 0.31	-0.24 ± 0.14	-0.69 ± 0.11

binning (light gray) often created a wide spread in $\text{RMSE}(\Delta\theta)$, while phase binning (black) mostly resulted in more concentrated data at $\text{RMSE}(\Delta\theta) \approx 2^\circ$. Since noise and streaking are strongly correlated with $\text{RMSE}(\Delta\theta)$, a wider spread of $\text{RMSE}(\Delta\theta)$ is reflected in larger RMSE values of SNR and CNR seen in Figs. 7(a) and 7(b), i.e., larger interbin image quality variation due to displacement-based binning.

In addition, it can be seen that equal-projection-density binning resulted in a more confined spread of $\text{mean}(\Delta\theta)$ values than displacement binning [see Fig. 9(b)], since the former aims for even allocation of projection images in all respiratory bins. However, even allocation of projection images did not lead to a more centralized distribution of $\text{RMSE}(\Delta\theta)$ values [see Fig. 9(c)], which was found to be the governing factor to image quality in Sec. 3.C.1. Consequently, similar interbin image quality variations were observed for both displacement and equal-projection-density displacement binning in Sec. 2.B.2.

3.C.3. Large maximum angular gap

Figure 9(a) shows that most of the $\text{max}(\Delta\theta)$ values clustered between 0° and 20° . Beyond this range, there were a few respiratory bins with much larger maximum angular gaps. These larger angular gaps resulted only from displacement binning (dark gray) and equal-projection-density displacement binning (light gray), but not phase binning (black). To demonstrate the effect of large maximum angular gaps on image quality, the reconstructed image with the largest maximum angular gap in this study, $\text{max}(\Delta\theta) = 78^\circ$ (the worst case), is displayed in Fig. 10. The large maximum angular gap resulted in insufficient projection coverage, causing major loss of image details which clearly degrades the clinical usability of displacement-based binning images.

It is worth pointing out that image quality metric calculations were still valid in these images because the region of interest for metric calculations was “reasonably” reconstructed such that the central homogenous region (SNR), fiducial marker (CNR), and the inner lung boundary (ERW) are still identifiable. Images with large maximum angular gaps resulted in considerably worse metric values than the other respiratory bins. However, large maximum angular gaps mostly

occurred only in the end inhalation and exhalation bins. This is because large baseline shifts or varying respiratory amplitudes are likely to cause the respiratory displacement to miss the end inhalation and exhalation displacement range, while there are in general at least one projection image allocated to each of the other bins in every respiratory cycle. Consequently, since large maximum angular gaps degraded the image quality of only one or two respiratory bins out of a total of ten in most cases, the mean SNR, CNR, and ERW values are not tremendously affected.

4. DISCUSSION

4.A. Respiratory signals

In this study, the three respiratory signal sources (RPM, intensity analysis, and marker tracking) showed observable differences in respiratory motion features (see Table I). It is noteworthy that although a direct comparison between respiratory signal sources requires some amplitude scaling, apparent disparities in motion estimation can still be observed in Fig. 1 (e.g., different baseline shift patterns). Previous studies have also shown observable mismatch between external and internal motion measurements.^{6,7} Despite the disparity in motion estimation between sources of respiratory signals, differences in the images were found to be negligible both qualitatively and quantitatively for the resolution used in this study ($1 \times 1 \times 2$ mm). Although a significant decrease in CNR value was observed in scan 2 with ASD-POCS reconstruction, it is most likely caused by occasional oversmoothing from the ASD-POCS algorithm. Based on the three-patient study (patients 1–3), the source of respiratory signals was concluded to have minimal impact on motion artifacts and interbin image quality variation. This confirms expectations since the division of motions into ten phases only provides a rough estimate of 4D anatomical information. Therefore, all reasonable estimations of respiratory motion should perform similarly well and should not have a significant impact on the use of 4D-CBCT in IGRT. Nonetheless, these conclusions need to be further tested in a larger patient cohort.

4.B. Respiratory binning

Displacement and equal-projection-density displacement binning (displacement-based binning) were shown to produce higher marker CNR values than phase binning in more than half of the 4D reconstruction cases. Since the marker CNR value is a surrogate for assessing possible motion blur in fine anatomical details (see Sec. 2.D.2), displacement-based binning is expected to result in more accurate and clearer representation of small moving structures (e.g., markers, vessels, or small lesions) based on the seven-patient study. A qualitative demonstration of this result is included in Appendix A. This concurs with results previously studied for 4D-CT (Refs. 17–19) in that displacement-based binning methods perform better in terms of capturing motion detail. It is noteworthy that conventional displacement binning performed slightly better than equal-projection-density displacement

FIG. 10. A FDK image reconstructed with marker signal and equal-projection-density displacement binning. This particular image corresponds to the right most light gray cross [EPDB (marker)] in Fig. 9(a) for patient 1, scan 2, and has a maximum angular gap of 78° , causing major loss of image details. This is the worst case example. The window level is set to $[0, 0.02] \text{ mm}^{-1}$ (attenuation coefficient).

binning in terms of motion details. This may be due to uneven divisions of displacement bins in equal-projection-density displacement binning.

Although more advantageous to the clinical use of 4D-CBCT than phase binning, displacement-based binning methods suffer from two major disadvantages—large maximum angular gaps and interbin image quality variation.

The occasional occurrence of large maximum angular gaps is due to the limited ability of displacement-based binning to handle baseline shifts. Large and long-lasting baseline shifts [e.g., marker signal of patient 1, scan 2 in Fig. 1(c)] or varying respiratory amplitudes can often result in the lack of projection images in the end inhalation and exhalation bins in a large time interval, therefore leading to a large maximum angular gap. This often results in very unreliable image quality in the end inhalation and exhalation bins as shown Fig. 10. Out of the two displacement-based binning methods, equal-projection-density binning was prone to larger maximum angular gap as can be seen in Fig. 9(a). This is a consequence of using the global maximum and minimum as the upper and lower boundaries (see Sec. 2.B.2.c), causing equal-projection-density displacement binning to miss more projection angular views in the end inhalation and exhalation bins. Conventional displacement binning handled baseline shifts with appropriate selections of the upper and lower displacement boundaries (see Sec. 2.B.2.b), therefore produced much more acceptable maximum angular gap values. In contrast, phase binning does not suffer from baseline shifts at all in terms of angular gaps, and appears to be the most reliable practical solution currently available. However, improved implementation of audiovisual biofeedback is expected to reduce baseline shifts considerably,³² making displacement-based binning a feasible option in the near future. Another possible solution is to eliminate large angular gaps by sharing certain projection views with neighboring respiratory bins.

Interbin image quality variation results from the nonlinear relationship between displacement and time in respiratory motions, which leads to uneven time intervals between displacement-based respiratory bins. Since projection images are acquired with a fixed frame rate, such uneven time intervals produce very different $\text{RMSE}(\Delta\theta)$ values between respiratory bins. The inconsistent $\text{RMSE}(\Delta\theta)$ values result in considerable variations in image quality between respiratory bins, which could lead to potential problems in the clinical implementation of displacement-based binning. Nevertheless, this may be resolved by an image acquisition technique such as respiratory motion guided-4DCBCT (RMG-4DCBCT),²⁷ where evenly spaced projection views are achieved by regulating the gantry velocity and imaging frequency. In this case, displacement-based binning could potentially replace phase binning as it is less prone to motion artifacts, as shown in this study.

4.C. Reconstruction algorithm

This study has comprehensively demonstrated that, both qualitatively and quantitatively, different reconstruction algorithms produce significant and consistent differences in im-

age quality. MKB and ASD-POCS, in particular, produce notable improvements in SNR and CNR values compared to FDK, implying that in practice, anatomical details are sharper, which may make them better choices for assessing margins in IGRT. The MKB algorithm only slightly outperformed FDK in terms of SNR and CNR, but gave noticeable qualitative improvements. On the other hand, ASD-POCS reconstructions produced the highest SNR and CNR values, but caused slight blurring in the bony anatomy (see Fig. 6) and lung boundary [higher ERW values in Figs. 7(c) and 8(b)]. Overall, both MKB and ASD-POCS improved the image quality considerably, but it is difficult to determine from the results presented in this study whether one outperforms the other. Nevertheless, MKB is considered the best practical solution overall due to the much smaller computational expense. This concurs with the study of Bergner *et al.*,²³ where MKB and ASD-POCS have been compared with several other reconstruction algorithms.

It is worth emphasizing that the increased CNR values from using the MKB or ASD-POCS algorithm are due to the reduction of noise and streaking, while the increased CNR values from using displacement-based binning are due to the reduction of motion artifacts. Since noise/streaking and motion artifacts are different types of degradation, a better reconstruction cannot compensate for the motion blur caused by phase binning. That is, in terms of improving the clarity of anatomical detail, a better reconstruction algorithm and respiratory binning method are both necessary. Nevertheless, the clinical usefulness of displacement-based binning is yet to be demonstrated as it suffers from large interbin image quality variation and large maximum angular gaps (see Sec. 4.B). Moreover, displacement-based binning is expected to benefit only about half of the 4D reconstruction cases [cf. Fig. 8(a)], while both MKB and ASD-POCS are likely to improve image quality significantly for all scans [cf. Fig. 8(b)]. At present, therefore, development in better reconstruction algorithms is considered a more effective strategy in improving 4D-CBCT image quality.

4.D. Projection angular spacing

$\text{RMSE}(\Delta\theta)$ was determined to be the most important projection angular spacing factor influencing image quality. Furthermore, Table III shows that SNR values were always strongly correlated with $\text{RMSE}(\Delta\theta)$ regardless of the reconstruction algorithm used, implying that reducing $\text{RMSE}(\Delta\theta)$ is as important as developing better reconstruction algorithms. Correlations were also observed between SNR and $\max(\Delta\theta)$. This is expected since a larger maximum angular gap is likely to result in a large $\text{RMSE}(\Delta\theta)$ in most cases.

Most of the maximum angular gaps were found to be clustered between 0° and 20° [cf. Fig. 9(a)]. Since $\max(\Delta\theta)$ was found to have a smaller impact on image quality than $\text{RMSE}(\Delta\theta)$, this indicates that a maximum angular gap well within this range may be acceptable in some cases (see Appendix C for an example). On the other hand, large maximum angular gaps ($\max(\Delta\theta) \gg 20^\circ$) can result in serious loss of image details (cf. Fig. 10), and need to be avoided. It has

been discussed in Sec. 4.B that large maximum angular gaps only occur when displacement-based binning is implemented on patients with large baseline shifts in the respiratory motions. This can be seen by comparing the distribution of maximum angular gaps for patient 1 and patient 3 in Fig. 9(a). Displacement-based binning clearly resulted in more large maximum angular gaps for patient 1 than for patient 3, the former shown to have a much more irregular respiratory motion than the latter in Fig. 1. Consequently, this study suggests that displacement-based binning methods should only be used for patients with relatively regular respiratory motions (e.g., patient 3).

Out of all three angular factors, $\text{mean}(\Delta\theta)$ is the least correlated with SNR. As $\text{mean}(\Delta\theta)$ represents the reciprocal of the number of projections, the result implies that the angular distribution of projections is more important than the number of projection images. This explains why no improvement in interbin image quality variation was observed with equal-projection-density displacement binning even though it aims for even allocation of projection images (see Sec. 3.C.2). This also justifies image acquisition techniques such as respiratory triggered-4DCBCT,²⁶ which reduces additional imaging dose that might otherwise cause projection clustering.

4.E. Limitations

An inevitable limitation of this study was the lack of ground truth for patient data, in which case the use of image quality metrics was necessary. Global metrics, such as the sum of absolute differences, are inappropriate for complicated patient images. Metrics based on local measurements are therefore more suitable for this type of study, but are not guaranteed to capture all global features. Nevertheless, a qualitative assessment was conducted to prevent any possible bias from using image quality metrics. A follow-up phantom study would provide ground truth and allow direct inter- and intrafactor comparisons.

Our study of respiratory signal sources was limited by the small number of available clinical scans (three patients, six clinical scans), because in order to obtain respiratory signal via fiducial marker tracking, only scans with at least one fiducial marker visible in all projection frames can be used, making most of the half-fan scans unusable. Consequently, the conclusion of respiratory signal sources having minimal impact on image quality is yet to be verified in a larger patient cohort. The results for the other factors were obtained exploiting all the available clinical data from patients with implanted markers in the cancer trial (seven patients), making the observations for respiratory binning methods, reconstruction algorithms, and projection angular spacing more statistically reliable relative to that of respiratory signal sources. Nevertheless, a future study involving a larger patient cohort may allow stronger conclusions to be made.

The images in this study were reconstructed with a resolution of $1 \times 1 \times 2$ mm. Although the source of respiratory signals was found to have no apparent impact on motion artifacts, it is yet to be verified for higher resolution cases. Nevertheless, the resolution of clinical 4D-CBCT images is

generally lower than $1 \times 1 \times 2$ mm. Therefore, the resolution used in this study is sufficiently relevant to the IGRT purpose of 4D-CBCT.

In this study, around 2400 projection images were used for each 4D reconstruction, whereas in practice this number is nearly halved. Future studies involving 1200 projection images for each 4D reconstruction may better represent the true clinical situation. However, a better choice of reconstruction algorithm is expected to be even more critical when the number of projections is not sufficient to reconstruct an image that can be used for IGRT.

5. CONCLUSION

This study is the first to quantitatively investigate the impact of different respiratory signal sources, respiratory binning methods, reconstruction algorithms, and projection angular spacing on thoracic 4D-CBCT image quality. Among the three controllable factors, respiratory signals appear to have no significant effect on image quality. Displacement-based binning was found to outperform phase binning in terms of capturing respiratory motion-related detail in more than half of the cases. Both MKB and ASD-POCS reconstruction algorithms significantly reduce noise and streaking, which are persistent artifacts in FDK reconstructions. Regarding projection angular spacing, acquisition of evenly distributed projection angular views was found to be more important than the number of projection images in each respiratory bin.

Overall, it has been shown that the use of displacement-based binning, choice of reconstruction algorithm, and the acquisition of evenly distributed projection angular views for every respiratory bin are the most important factors to consider for improving the quality of thoracic 4D-CBCT images for application in IGRT. However, displacement-based binning is prone to large interbin image quality variation and occasional large angular gaps, making its practical implementation challenging. Moreover, it was found to benefit only about half of the cases studied, while better reconstruction algorithms were shown to improve image quality almost 100% of the time. Projection angular spacing is not directly controllable with current technology. It is therefore suggested that improving reconstruction algorithms be the most effective strategy for improving thoracic 4D-CBCT image quality.

ACKNOWLEDGMENTS

The authors would like to thank Assistant Professor Geoffrey Hugo and Dr. Simon Rit for providing guidance in using Reconstruction Toolkit. This project is supported by an NHMRC Australia Fellowship, NHMRC project Grant No. 1034060, US NCI Grant No. P01CA116602, an International Postgraduate Research Scholarships, and an Australian Postgraduate Award.

APPENDIX A: PHASE BINNING VS DISPLACEMENT BINNING

To demonstrate the superior performance of displacement binning in relation to motion detail, images of the fiducial marker in all ten respiratory bins (patient 3, scan 5) reconstructed using phase binning and displacement binning are displayed in Figs. 11(a) and 11(b). Apparent motion artifacts are observed for phase binning, especially in the 50%, 60%, 70%, and 80% frames. The shape of the marker appears blurred by motion artifacts, resulting in indistinct motion trajectory. In contrast, there is no evidence of motion blur in the displacement binning images. Clear respiratory motion can also be seen from the marker trajectory.

APPENDIX B: FDK VS ASD-POCS

Figures 11(b) and 11(c) compare the displacement binning images reconstructed using FDK and ASD-POCS for all ten respiratory bins. It can be seen that ASD-POCS reduced noise and streaking without causing observable blur to the fiducial marker. As a result, the motion of the marker is observed more clearly.

APPENDIX C: MAXIMUM ANGULAR GAP NEAR 20°

Figure 12 displays FDK reconstructed images with a maximum angular gap of 8.8° and 18.5°. Apart from the apparent streaking near the right lung region for the $\max(\Delta\theta) = 18.5^\circ$ image, the rest of the image was reasonably reconstructed. Most of the anatomical details can still be seen clearly. This suggests that a maximum angular gap below approximately 20° may be acceptable in some cases for the clinical implementation of 4D-CBCT.

FIG. 12. The axial views of FDK reconstructed images with a maximum angular gap of 8.8° (left) and 18.5° (right) from patient 1, scan 1. The window level was set to $[0, 0.02] \text{ mm}^{-1}$ (attenuation coefficient).

APPENDIX D: STUDENT'S *t*-TESTS

Two-tailed paired Student's *t*-tests were employed to determine (i) whether the alternative respiratory binning methods (DB, EPDB) performed similarly, better, or worse compared to the current clinical respiratory binning method (PB) in terms of SNR, CNR, and ERW; (ii) whether the alternative reconstruction algorithms (MKB, ASD-POCS) performed similarly, better, or worse compared to the current clinical reconstruction algorithm (FDK) in terms of SNR, CNR, and ERW. The flow chart of the paired *t*-test analysis is displayed in Fig. 13, and also explained below. The percentages of cases in which a particular alternative practice performed current clinical practice were calculated and shown in Fig. 8.

Each DB or EPDB 4D reconstruction using a particular respiratory signal source and reconstruction algorithm was compared to the corresponding PB 4D reconstruction using a two-tailed paired Student's *t*-test with the null hypothesis that DB/EPDB and PB produce similar metric values (SNR, CNR, ERW). In other words, each paired *t*-test consists of

FIG. 11. The coronal views of the fiducial marker in all ten respiratory bins from patient 3, scan 5, reconstructed using: (a) PB and FDK algorithm; (b) DB and FDK algorithm; and (c) DB and ASD-POCS algorithm. All the images were reconstructed using the marker signal. The window level was set to $[0, 0.04] \text{ mm}^{-1}$ (attenuation coefficient). Two dashed lines were plotted to align with the top and bottom most positions of the marker in its motion trajectory in the displacement binning (FDK) case. The same dashed lines were plotted for phase binning (FDK) and displacement binning (ASD-POCS) to assist comparison between the three cases.

FIG. 13. A flow chart illustrating how the two-tailed paired *t*-tests were employed to determine whether a particular alternative practice (DB, EPDB; MKB, ASD-POCS) performed similarly, better, or worse compared to the corresponding current clinical practice (PB; FDK).

two data series (one for DB/EPDB, one for PB), each series containing ten metric values corresponding to the ten respiratory bins. DB or EPDB performed similar to PB if the *p*-value was greater or equal to 0.05. DB or EPDB performed better/worse than PB if the *p*-value was smaller than 0.05, and produced higher/lower mean SNR, higher/lower mean CNR, or lower/higher mean ERW values than PB. Six DB-PB and six EPDB-PB paired *t*-tests ([RPM, marker] $\times$ [FDK, MKB, ASD-POCS]) were conducted in each of scans 1–6; three paired *t*-tests ([RPM] $\times$ [FDK, MKB, ASD-POCS]) were conducted for each of scans 7–14. Intensity analysis was not included in the *t*-tests since it can only be used with phase binning. This makes a total of $6 \times 6 + 8 \times 3 = 60$ paired *t*-tests each for DB-PB and EPDB-PB comparisons. The percentages in Fig. 8(a) represent the number of cases out of 60.

Similarly, each MKB or ASD-POCS 4D reconstruction was compared to the corresponding FDK 4D reconstruction using the same two-tailed paired Student's *t*-test. The definition of “similar/better/worse” performance of MKB or ASD-POCS relative to FDK was similar to that described earlier. Seven MKB-FDK and seven ASDPOCS-FDK paired *t*-tests ([RPM, marker] $\times$ [PB, DB, EPDB] + [IA] $\times$ [PB]) were conducted in each of scans 1–6; three paired *t*-tests ([RPM] $\times$ [PB, DB, EPDB]) were conducted for each of

scans 7–14. This makes a total of $6 \times 7 + 8 \times 3 = 66$ paired *t*-tests each for MKB-FDK and ASDPOCS-FDK comparisons. The percentages in Fig. 8(b) represent the number of cases out of 66.

^{a)} Author to whom correspondence should be addressed. Electronic mail: paul.keall@sydney.edu.au; Telephone: +61 2 9351 3590.

¹ J.-J. Sonke, L. Zijp, P. Remeijer, and M. van Herk, “Respiratory correlated cone beam CT,” *Med. Phys.* **32**(4), 1176–1186 (2005).

² D. P. Gierga, J. Brewer, G. C. Sharp, M. Betke, C. G. Willett, and G. T. Y. Chen, “The correlation between internal and external markers for abdominal tumors: Implications for respiratory gating,” *Int. J. Radiat. Oncol., Biol., Phys.* **61**(5), 1551–1558 (2005).

³ Y. Seppenwoolde *et al.*, “Precise and real-time measurement of 3D tumor motion in lung due to breathing and heartbeat, measured during radiotherapy,” *Int. J. Radiat. Oncol., Biol., Phys.* **53**(4), 822–834 (2002).

⁴ P. R. Poulsen *et al.*, “A method for robust segmentation of arbitrarily shaped radiopaque structures in cone-beam CT projections,” *Med. Phys.* **38**(4), 2151–2156 (2011).

⁵ C. Nelson, G. Starkschall, P. Balter, R. C. Morice, C. W. Stevens, and J. Y. Chang, “Assessment of lung tumor motion and setup uncertainties using implanted fiducials,” *Int. J. Radiat. Oncol., Biol., Phys.* **67**(3), 915–923 (2007).

⁶ J. D. P. Hoisak, K. E. Sixel, R. Tirona, P. C. F. Cheung, and J. P. Pignol, “Correlation of lung tumor motion with external surrogate indicators of respiration,” *Int. J. Radiat. Oncol., Biol., Phys.* **60**(4), 1298–1306 (2004).

⁷ H. Yan, F. F. Yin, G. P. Zhu, M. Ajlouni, and J. H. Kim, “The correlation evaluation of a tumor tracking system using multiple external markers,” *Med. Phys.* **33**(11), 4073–4084 (2006).

⁸ R. I. Berbeco, H. Mostafavi, G. C. Sharp, and S. B. Jiang, “Towards fluoroscopic respiratory gating for lung tumours without radiopaque markers,” *Phys. Med. Biol.* **50**(19), 4481–4490 (2005).

⁹ A. Kavanagh, P. M. Evans, V. N. Hansen, and S. Webb, “Obtaining breathing patterns from any sequential thoracic x-ray image set,” *Phys. Med. Biol.* **54**(16), 4879–4888 (2009).

¹⁰ L. Zijp, J. J. Sonke, and M. van Herk, “Extraction of the respiratory signal from sequential thorax cone-beam X-ray images,” in *Proceedings of ICCR*, Seoul, Korea (Jeong Publishing, Seoul, 2004), pp. 507–509.

¹¹ M. van Herk, L. Zijp, P. Remeijer, J. Wolthaus, and J. J. Sonke, “On-line 4D cone beam CT for daily correction of lung tumor position during hypofractionated radiotherapy,” in *Proceedings of ICCR*, Toronto, Canada (Novel Digital Publishing, Toronto, 2007), p. 6241.

¹² I. Vergalasova, J. Cai, and F. F. Yin, “A novel technique for markerless, self-sorted 4D-CBCT: Feasibility study,” *Med. Phys.* **39**(3), 1442–1451 (2012).

¹³ H. Yan *et al.*, “Extracting respiratory signals from thoracic cone beam CT projections,” *Phys. Med. Biol.* **58**(5), 1447–1464 (2013).

¹⁴ S. A. Nehmeh *et al.*, “Quantitation of respiratory motion during 4D-PET/CT acquisition,” *Med. Phys.* **31**(6), 1333–1338 (2004).

¹⁵ V. R. Kini, S. S. Vedam, P. J. Keall, S. Patil, C. Chen, and R. Mohan, “Patient training in respiratory-gated radiotherapy,” *Med. Dosim.* **28**(1), 7–11 (2003).

¹⁶ E. Yorke, K. E. Rosenzweig, R. Wagman, and G. S. Mageras, “Interfractional anatomic variation in patients treated with respiration-gated radiotherapy,” *J. Appl. Clin. Med. Phys.* **6**(2), 19–32 (2005).

¹⁷ M. J. Fitzpatrick *et al.*, “Displacement-based binning of time-dependent computed tomography image data sets,” *Med. Phys.* **33**(1), 235–246 (2006).

¹⁸ A. F. Abdelnour *et al.*, “Phase and amplitude binning for 4D-CT imaging,” *Phys. Med. Biol.* **52**(12), 3515–3529 (2007).

¹⁹ H. Li *et al.*, “Clinical evaluations of an amplitude-based binning algorithm for 4DCT reconstruction in radiation therapy,” *Med. Phys.* **39**(2), 922–932 (2012).

²⁰ L. A. Feldkamp, L. C. Davis, and J. W. Kress, “Practical cone-beam algorithm,” *J. Opt. Soc. Am. A* **1**(6), 612–619 (1984).

²¹ G. C. McKinnon and R. H. T. Bates, “Towards imaging the beating heart usefully with a conventional CT scanner,” *IEEE Trans. Biomed. Eng.* **BME-28**(2), 123–127 (1981).

²² S. Leng *et al.*, “Streaking artifacts reduction in four-dimensional cone-beam computed tomography,” *Med. Phys.* **35**(10), 4649–4659 (2008).

²³ F. Bergner *et al.*, “An investigation of 4D cone-beam CT algorithms for slowly rotating scanners,” *Med. Phys.* **37**(9), 5044–5053 (2010).

- ²⁴E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory* **52**(2), 489–509 (2006).
- ²⁵E. Y. Sidky and X. C. Pan, "Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization," *Phys. Med. Biol.* **53**(17), 4777–4807 (2008).
- ²⁶B. J. Cooper, R. T. O'Brien, S. Balik, G. D. Hugo, and P. J. Keall, "Respiratory triggered 4D cone-beam computed tomography: A novel method to reduce imaging dose," *Med. Phys.* **40**(4), 041901 (9pp.) (2013).
- ²⁷R. T. O'Brien, B. J. Cooper, and P. J. Keall, "Optimizing 4D cone beam computed tomography acquisition by varying the gantry velocity and projection time interval," *Phys. Med. Biol.* **58**(6), 1705–1723 (2013).
- ²⁸N. O. Roman, W. Shepherd, N. Mukhopadhyay, G. D. Hugo, and E. Weiss, "Interfractional positional variability of fiducial markers and primary tumors in locally advanced non-small-cell lung cancer during audiovisual biofeedback radiotherapy," *Int. J. Radiat. Oncol., Biol., Phys.* **83**(5), 1566–1572 (2012).
- ²⁹R. J. Cropp, "Implementation of respiratory-correlated cone-beam CT on Varian linac systems," Master thesis (The University Of British Columbia, 2011), p. 48.
- ³⁰G. S. L. Zeng and G. T. Gullberg, "Unmatched projector/backprojector pairs in an iterative reconstruction algorithm," *IEEE Trans. Med. Imaging* **19**(5), 548–555 (2000).
- ³¹S. Rit, M. Vila Oliva, S. Brousmiche, R. Labarbe, D. Sarrut, and G. C. Sharp, "The reconstruction toolkit (RTK), an open-source cone-beam CT reconstruction toolkit based on the Insight Toolkit (ITK)," in *International Conference on the Use of Computers in Radiation Therapy (ICCR)*, Melbourne, Australia, 2013.
- ³²S. Pollock, D. Lee, P. Keall, and T. Kim, "Audiovisual biofeedback improves motion prediction accuracy," *Med. Phys.* **40**(4), 041705 (9pp.) (2013).