Effects of calibration methods on quantitative material decomposition in photon-counting spectral computed tomography using a maximum *a posteriori* estimator

Tyler E. Curtis and Ryan K. Roeder[a)]
Department of Aerospace and Mechanical Engineering, Bioengineering Graduate Program, University of Notre Dame, Notre Dame, IN 46556, USA

(Received 23 February 2017; revised 27 June 2017; accepted for publication 30 June 2017; published 8 August 2017)

Purpose: Advances in photon-counting detectors have enabled quantitative material decomposition using multi-energy or spectral computed tomography (CT). Supervised methods for material decomposition utilize an estimated attenuation for each material of interest at each photon energy level, which must be calibrated based upon calculated or measured values for known compositions. Measurements using a calibration phantom can advantageously account for system-specific noise, but the effect of calibration methods on the material basis matrix and subsequent quantitative material decomposition has not been experimentally investigated. Therefore, the objective of this study was to investigate the influence of the range and number of contrast agent concentrations within a modular calibration phantom on the accuracy of quantitative material decomposition in the image domain.
Methods: Gadolinium was chosen as a model contrast agent in imaging phantoms, which also contained bone tissue and water as negative controls. The maximum gadolinium concentration (30, 60, and 90 mM) and total number of concentrations (2, 4, and 7) were independently varied to systematically investigate effects of the material basis matrix and scaling factor calibration on the quantitative (root mean squared error, RMSE) and spatial (sensitivity and specificity) accuracy of material decomposition. Images of calibration and sample phantoms were acquired using a commercially available photon-counting spectral micro-CT system with five energy bins selected to normalize photon counts and leverage the contrast agent *k*-edge. Material decomposition of gadolinium, calcium, and water was performed for each calibration method using a maximum *a posteriori* estimator.
Results: Both the quantitative and spatial accuracy of material decomposition were most improved by using an increased maximum gadolinium concentration (range) in the basis matrix calibration; the effects of using a greater number of concentrations were relatively small in magnitude by comparison. The material basis matrix calibration was more sensitive to changes in the calibration methods than the scaling factor calibration. The material basis matrix calibration significantly influenced both the quantitative and spatial accuracy of material decomposition, while the scaling factor calibration influenced quantitative but not spatial accuracy. Importantly, the median RMSE of material decomposition was as low as ~1.5 mM (~0.24 mg/mL gadolinium), which was similar in magnitude to that measured by optical spectroscopy on the same samples.
Conclusion: The accuracy of quantitative material decomposition in photon-counting spectral CT was significantly influenced by calibration methods which must therefore be carefully considered for the intended diagnostic imaging application.
[https://doi.org/10.1002/mp.12457]

Key words: calibration phantom, contrast media, material decomposition, photon counting detector, spectral computed tomography

1. INTRODUCTION

X-ray computed tomography (CT) provides low cost anatomic and diagnostic imaging at high spatial and temporal resolution.[1–3] Conventional clinical and laboratory CT imaging systems utilize a polychromatic x-ray source with a photon energy spectrum ranging from a lower threshold of ~20 keV, set by beam filtration, to a peak tube potential of 80 to 140 kVp.[3] Materials and tissues exhibit photon energy-dependent differences in x-ray attenuation coefficients over this energy range.[4,5] However, conventional CT systems utilize an energy integrating detector, which sums the photon counts over the entire energy spectrum, ignoring energy-specific spectral information.[6] Dual-energy CT (DECT) captures limited energy-dependent differences by imaging at two peak tube potentials and applying a weighted subtraction to decompose and quantify material composition.[7,8] However, DECT increases radiation exposure and is typically limited to the separation of two materials, usually soft tissue and hard attenuators, such as bone or contrast media.[7,8] Multi-energy acquisition is required to decompose a greater number of material and tissue compositions.

Recent advances in energy-sensitive x-ray detectors have enabled multi-energy or spectral CT using a single polychromatic x-ray source.[9,10] Energy-sensitive detectors measure the number of photon interactions with the detector and the energy of each interaction by the charge output of the detector chip. Advances in photon-counting detectors have decreased the computational time required to evaluate these interactions, allowing photon energies to be accurately measured in high flux, multipixel applications.[11,12] Photon-counting detectors, such as the Medipix 3RX, make count corrections for charge summing and pulse pileup, based upon the pulse height and analysis of coinciding photon interactions, decreasing spectral blurring and increasing image quality.[11,12] The measured energy-dependent attenuation coefficients enable identification of material composition (material decomposition).[10,13–17] Material decomposition has been investigated with photon-counting spectral CT using both simulated[10,15–18] and acquired[14,18,19] data sets, and these investigations have primarily focused on algorithm design.

Spectral unmixing or material decomposition can be accomplished in the projection[10,13,16–21] or image[14,15] domain using, for example, least squares regression models,[15] maximum likelihood estimators,[10,18,20] or maximum *a posteriori* (MAP) estimators.[21] Supervised decomposition methods utilize an estimated attenuation for each material of interest within each energy bin. The estimated attenuation must be calibrated based upon calculated or measured values for known compositions. Measurements using a calibration phantom are advantageous in accounting for system-specific sources of variability which are not readily described by tabulated or simulated data. However, the effect of calibration methods on the material basis matrix and subsequent quantitative material decomposition has not been experimentally investigated.

Therefore, the objective of this study was to investigate the influence of the range and number of contrast agent concentrations within a modular calibration phantom on the accuracy of quantitative material decomposition in the image domain. A systematic analysis of calibration phantom concentrations was performed using a commercially available photon-counting spectral micro-CT system with five energy bins selected to normalize photon counts and leverage the contrast agent *k*-edge. Material basis matrix values were estimated using multiple linear regression models. An increase in the maximum concentration (range) and number of concentrations utilized in the calibration was hypothesized to result in an increased signal-to-noise ratio in each energy bin for improved accuracy and precision in the material basis matrix values, and thus improved accuracy of quantitative material decomposition.

2. METHODS

2.A. Calibration phantom

Gadolinium was chosen as a model contrast agent for spectral CT due to clinical use as a contrast agent for diagnostic imaging and the presence of a *k*-shell absorption edge within the clinical x-ray energy spectrum. Gadolinium (III) nitrate hexahydrate, $Gd(NO_3)_3{\cdot}6H_2O$ (Acros Organics, 99.9%), solutions were prepared in deionized (DI) water at selected concentrations ranging from 0 to 90 mM through dilutions from a common stock solution prepared at 100 mM. The maximum gadolinium concentration (30, 60, and 90 mM) and total number of concentrations (2, 4, and 7) were independently varied to investigate the effect on quantitative material decomposition (Table I). Calcium signal was calibrated using a custom bone mineral density phantom with compositions ranging from 0 to 60 vol% (0 to 1860 mg/cm^3) hydroxyapatite in polyethylene.[22]

Gadolinium concentrations were verified by inductively coupled plasma optical emission spectroscopy (ICP-OES, Optima 7000, Perkin Elmer, Waltham, MA, USA) after digesting samples in 5% nitric acid. Calibration curves for ICP-OES were created by diluting certified standard gadolinium solutions (SPEX CertiPrep, assurance grade, Metuchen, NJ, USA). Measured gadolinium concentrations were compared with the expected concentrations using linear least squares regression and a paired *t*-test (JMP® 11.0, SAS Institute, Inc., Cary, NC, USA).

2.B. Sample phantom

A sample phantom containing 0, 5, and 10 mM concentrations of gadolinium nitrate dissolved in DI water, as well as a rabbit femur embedded in 1% agarose (Thermo Scientific, molecular biology grade), was used to evaluate the material decomposition calibrations. The range of contrast agent concentrations was selected to be feasible for *in vivo* targeted delivery to a site of interest, based upon preclinical animal models, while challenging the detection limit of conventional CT.[23]

2.C. Image acquisition

For both the calibration and sample phantoms, Eppendorf tubes containing each gadolinium and calcium concentration were placed in a modular phantom for imaging. Images of both the calibration and sample phantoms were acquired using a commercially available spectral CT (MARS-12 v.5, MARS Bioimaging Ltd., Christchurch, New Zealand)

TABLE I. The range and number of gadolinium contrast agent concentrations used to systematically investigate the effect of calibration methods on quantitative material decomposition.

Concentration range	Maximum concentration (mM)	Number of concentrations	Step size (mM)
30 × 2	30	2	30
30 × 4	30	4	10
30 × 7	30	7	5
60 × 2	60	2	60
60 × 4	60	4	20
60 × 7	60	7	10
90 × 2	90	2	90
90 × 4	90	4	30
90 × 7	90	7	15

equipped with a polychromatic x-ray source operating at 120 kVp, 1.96 mm aluminum beam filtration, and a photon-counting detector comprising a CdZnTe semiconductor sensor bonded to a Medipix 3RX chip with five energy bins. Energy thresholds were set at 30.9, 50.0, 60.1, and 73.1 keV to normalize photon counts across energy bins and leverage the *k*-edge discontinuity of gadolinium at 50.2 keV.[24] Each pixel was limited to a mean of 2000 counts total across all energy bins to mitigate the effects of pulse pileup. Reconstructions were performed with a 100 μm isometric voxel size and a nominal resolution of ~300 μm. Five replicates of the calibration and sample phantoms were imaged for each calibration method (Table I). Images for all calibration and sample phantoms were acquired over four separate days of imaging with a flatfield correction image taken each day. The phantom design and relatively low contrast agent concentrations obviated the need for beam hardening correction due to the absence of hard attenuators in-line with the x-ray source. The measured x-ray attenuation was correlated with gadolinium concentration for each energy bin using linear least squares regression (JMP® 11.0). For representative images, grayscale values were converted to Hounsfield units (HU) by calibration with air (−1000 HU) and water (0 HU).

2.D. Material basis matrix

The material basis matrix, C, is an $M \times N$ matrix of the estimated x-ray attenuation for each material in each energy bin, where M is the number of energy bins and N is the number of materials to be decomposed. Attenuation is assumed to be approximately linear for each composition on a bin-by-bin basis. The measured effective attenuation within the jth voxel for the kth energy bin can be described by,

$$\bar{\mu}_{jk} = \sum_{i=1}^{N} (\mu_{ijk} x_i) \tag{1}$$

where N is the total number of materials to be identified, μ_i is the mass attenuation coefficient of the ith material, and x_i is the volume fraction (or fractional abundance) of the ith material. The estimated attenuation of gadolinium in each energy bin was determined using a multiple linear regression of known gadolinium concentrations as,

$$\bar{\mu} = b_o + b_1 x_1 + b_2 x_2 \tag{2}$$

where $\bar{\mu}$ is the effective attenuation of voxels within the measured volume of interest (VOI), b_o is the intercept, which was assumed to be zero, b_1 is the estimated attenuation of gadolinium, and b_2 is the estimated attenuation of water. Given the known concentrations of gadolinium in each calibration range (Table I), the theoretical volume fraction of gadolinium and water in each composition was calculated using,

$$v_1 \approx \frac{(m_1/\rho_1)}{(m_1/\rho_1) + (m_2/\rho_2)} \tag{3}$$

where v_1 is the volume fraction of gadolinium, m_1 and m_2 are the mass of gadolinium and water respectively, ρ_1 and ρ_2 are the density of gadolinium and water, respectively, and the volume fraction of water is $v_2 \approx 1 - v_1$. The linear model was then computed using the estimated volume fraction of gadolinium and water, and the mean x-ray attenuation measured within a 97 mm^3 cuboidal VOI, for each calibration composition and imaging replicate. The calculation was repeated for each of the five energy bins to complete the basis matrix (C). The measured x-ray attenuations of gadolinium and water were paired for each calibration method and imaging replicate. The same process was repeated for the calcium calibration phantom, using the volume fraction of hydroxyapatite and polyethylene as the input variables. A single vector of estimated attenuation for calcium was calculated and used for all material basis matrices.

2.E. Material decomposition

Material decomposition of gadolinium, calcium, and water was performed on sample phantom images using a quadratic programming function (quadprog, MATLAB, v.9.0, MathWorks, Inc., Natick, MA, USA) with a MAP estimator[25] and material basis matrices calculated for each calibration method. A MAP estimator was chosen to mitigate the large number of possible solutions to the linear system of equations, which were constrained to satisfy both full additivity and non-negativity. The volume fraction (or fractional abundance) of materials in each voxel of sample phantom images was estimated by minimizing the equation,

$$min_x\, 0.5 \cdot x^T \cdot H \cdot x + f^T \cdot x \tag{4}$$

where x is a vector of volume fractions of length N, and

$$H = 2 \cdot C^T \cdot C \tag{5}$$

$$f = -2 \cdot C^T \cdot x \tag{6}$$

such that,

$$A \cdot x \leq b \tag{7}$$

$$A_{eq} \cdot x = b_{eq} \tag{8}$$

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & 1_N \\ -1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & -1_N \end{bmatrix} \tag{9}$$

where b is an $N \times 2$ vector containing N ones and N zeros, A_{eq} is a vector of length N comprised of ones, and b_{eq} is equal to one. Eq. (7) limits the solution to remain between 0 and 1, satisfying the non-negativity constraint. Eq. (8) ensures that the vector of volume fractions sums to 1 for each voxel, satisfying the full additivity constraint.

2.F. Postdecomposition scaling

Postdecomposition scaling was utilized to quantify the material decomposition as material concentration. The mean

x-ray attenuation of each concentration within the calibration VOI was determined for all energy bins. These mean attenuation values were then decomposed using the MAP estimator for each of the basis matrices used for material decomposition. The estimated volume fraction of each concentration was correlated with the known concentration (mM) in the calibration VOI. Linear least squares regression was performed for each possible combination of calibration methods (Table I) for both the material basis matrix and scaling factor. The intercept, b_o, was allowed to be nonzero to determine the best-fit regression line. Decomposed images were then scaled to millimolar (mM) concentrations using the linear regression resulting in a total of nine decompositions, each with nine distinct scaling factors, using the nine calibration methods (Table I).

2.G. Evaluation of quantitative material decomposition

Material decomposition was evaluated on sample phantom images for each permutation of the material basis matrix and scaling factor calibration (Table I). Evaluation of the effects of the calibration method on quantitative material decomposition was limited to the gadolinium concentration to focus on contrast media and simplify interpretation. True positive VOIs included the two concentrations of gadolinium; true negative VOIs included the water and rabbit bone. Each VOI was 101 mm^3 including 20 slices of the reconstruction within the inner diameter of each Eppendorf tube.

The accuracy of the quantitative material decomposition for each calibration method was evaluated by measuring the root mean squared error (RMSE) of the measured gadolinium concentrations compared with the known concentrations within the sample phantom. The spatial accuracy of the quantitative material decomposition for each calibration method was evaluated by measuring the specificity and sensitivity at a 5 mM threshold. Area under the receiver operating characteristics (ROC) curve (AUC) was calculated by evaluating the specificity and sensitivity over a range of concentrations (0–20 mM) which exceeded the concentrations within the sample phantom. The RMSE, specificity, sensitivity, and AUC for each possible combination of calibration methods for both the material basis matrix and scaling factor were compared using box plots showing the 10th, 25th, 50th (median), 75th, and 90th percentiles from the five replicates. Effects of the range and number of contrast agent concentrations in the material basis matrix and/or scaling factor calibration on RMSE, specificity, sensitivity, and AUC were examined using multiple and two-way analysis of variance (ANOVA) (JMP® 11.0). The level of significance for all tests was set at $P < 0.05$.

3. RESULTS

3.A. Calibration phantoms

The accuracy of gadolinium concentrations in the calibration phantom was verified by ICP-OES (Fig. 1). Measured gadolinium concentrations exhibited a strong linear correlation with the expected concentration ($P < 0.0001$, $R^2 = 0.999$, RMSE = 1.03 mM). Overall, measured gadolinium concentrations were greater than the expected concentrations ($P < 0.0001$, paired t-test) with a mean error of 3.2 mM and a RMSE of 3.4 mM (Table II). The RMSE of gadolinium concentrations in the phantom decreased slightly with a decreased maximum concentration or increased number of concentrations (Table II).

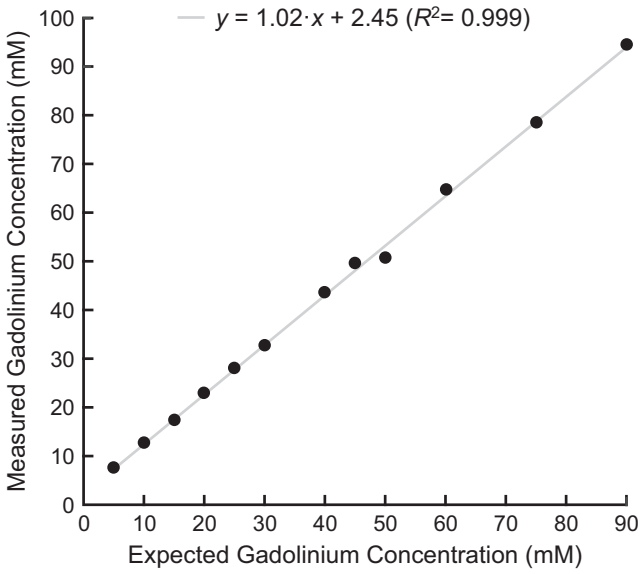

FIG. 1. Expected gadolinium concentrations in the calibration phantom compared with measurements by ICP-OES. Measured gadolinium concentrations exhibited a strong linear correlation with the expected concentration ($P < 0.0001$, $R^2 = 0.999$, RMSE = 1.03 mM), and were greater than expected concentrations overall ($P < 0.0001$, paired t-test) with a mean error of 3.2 mM and a RMSE of 3.4 mM (Table II). Note that the range of concentrations included and exceeded concentrations expected to be utilized in preclinical and clinical imaging.

Grayscale CT images exhibited increased x-ray attenuation with increased gadolinium concentration (Figs. 2 and 3), as expected, and no apparent beam hardening artifacts [Fig. 2(a)]. The maximum x-ray attenuation for a given gadolinium concentration occurred in the 50–60.1 keV energy bin (Figs. 2 and 3), corresponding to the k-edge of gadolinium. The measured x-ray attenuation exhibited a strong linear correlation ($R^2 > 0.95$) with gadolinium concentration for each energy bin of the photon-counting detector (Fig. 3). Therefore, material basis matrix values were estimated using multiple linear regression models for each calibration method (Table I). RMSE in multiple linear regression models was increased for higher energy bins, and with an increased maximum concentration (range) or number of concentrations in the calibration, but was generally less than 30 HU (Table II).

3.B. Overall effects of calibration methods on material decomposition

The overall effects of calibration methods on the quantitative (RMSE) and spatial (sensitivity, specificity, and AUC)

TABLE II. RMSE of linear regressions for gadolinium concentrations measured by ICP-OES versus expected concentrations, and x-ray attenuation measured by spectral CT for each energy bin of the photon-counting detector versus expected gadolinium concentrations, for each calibration method used to estimate the x-ray attenuation in the material basis matrix.

		x-ray attenuation (HU)				
Concentration range	ICP-OES (mM)	20.0–30.9 keV	30.9–50.0 keV	50.0–60.1 keV	60.1–73.1 keV	73.1–120 keV
30 × 2	2.9	3	3	4	10	16
30 × 4	2.9	4	3	6	11	16
30 × 7	2.8	9	6	9	12	16
60 × 2	4.8	5	5	6	20	40
60 × 4	3.9	8	7	7	17	33
60 × 7	3.2	16	10	15	17	31
90 × 2	4.5	3	4	5	9	20
90 × 4	4.2	16	12	12	34	45
90 × 7	3.9	17	12	13	30	38
Overall	3.4	15	10	23	25	34

accuracy of material decomposition were evaluated by varying the maximum concentration (30, 60, and 90 mM) and number of equally distributed concentrations (2, 4, and 7) within the calibration phantom (Table I) for both the material basis matrix and scaling factor calibrations. The maximum concentration used in the material basis matrix calibration significantly influenced the RMSE, specificity, sensitivity, and AUC ($P < 0.0001$, MANOVA) of the resulting material decomposition (Table III). Similarly, the maximum concentration used in the scaling factor calibration significantly influenced the RMSE, specificity, and sensitivity ($P < 0.001$), but not AUC. The number of concentrations used in the material basis matrix calibration significantly influenced the specificity and AUC ($P < 0.005$), but not RMSE and sensitivity, of the resulting material decomposition (Table III). In contrast, the number of concentrations

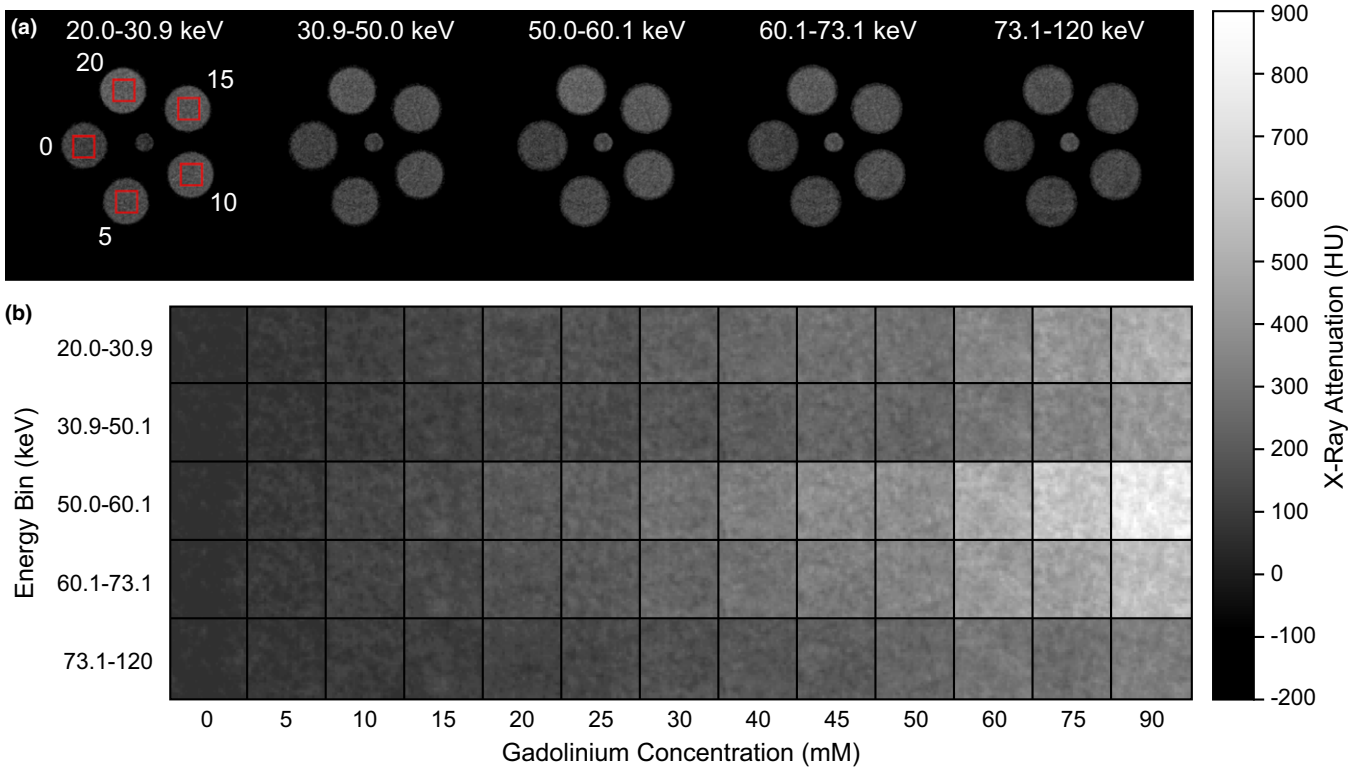

FIG. 2. Representative (a) grayscale CT image slices of a calibration phantom comprising of 0, 5, 10, 15, and 20 mM gadolinium concentrations showing example VOIs (boxes) used for determining material basis matrix calibrations, and (b) cropped grayscale images of VOIs spanning the full range of gadolinium concentrations, for each energy bin of the photon-counting detector. Note that all grayscale intensities were converted to HU. [Color figure can be viewed at wileyonlinelibrary.com]

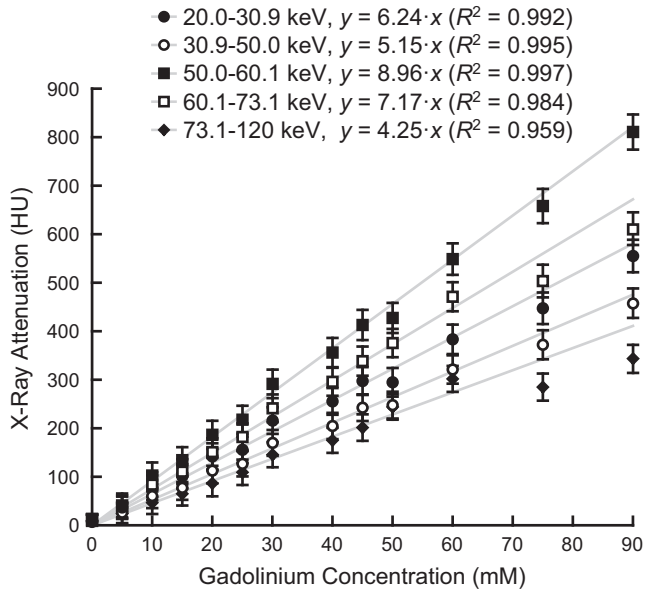

FIG. 3. The measured x-ray attenuation (HU) of all gadolinium concentrations for each energy bin of the photon-counting detector. Error bars show one standard deviation of the mean. Gray lines show the overall best fit line for each energy bin using linear least squares regression. The RMSE of linear regressions for each calibration method and each energy bin of the photon-counting detector is tabulated in Table II.

used in the scaling factor calibration significantly influenced the RMSE and sensitivity ($P < 0.05$), but not specificity and AUC. Significant interactions between the range and number of contrast agent concentrations in the material basis matrix calibration were observed for RMSE, specificity and AUC ($P < 0.001$), but not sensitivity. Significant interactions between the range and number of contrast agents in the scaling factor calibration were observed for RMSE, sensitivity, and specificity ($P < 0.001$), but not AUC.

3.C. Material basis matrix calibration

The effects of the calibration method on the material basis matrix calibration and subsequent material decomposition were further evaluated using the same calibration method (Table I) for both the material basis matrix and scaling factor calibrations. Representative gadolinium segmented images of the sample phantom after material decomposition qualitatively demonstrated that overall image quality was visually improved with an increased maximum concentration or number of concentrations in the calibration (Fig. 4). Specificity for gadolinium compared with hard attenuators was noticeably compromised using the lowest maximum concentration (30 mM) in the calibration as evidenced by the presence of gadolinium signal in the bone sample. On the other hand, specificity for gadolinium compared with soft attenuators (e.g., polypropylene sample tubes) that were not included in the basis matrix was improved using the lowest maximum concentration (30 mM) in the calibration. Specificity for gadolinium was visually improved for both hard and soft attenuators using an increased number of concentrations in the calibration (Fig. 4).

Quantitative evaluation of material decomposition revealed that calibration using a greater concentration range significantly improved the RMSE [Fig. 5(a)], specificity [Fig. 5(c)], and AUC [Fig. 5(d)] ($P < 0.0001$, ANOVA), but decreased sensitivity [Fig. 5(b)] ($P < 0.0005$, ANOVA). The effects of using a greater number of concentrations in the calibration on the RMSE, sensitivity, specificity, and AUC of the material decomposition were smaller in magnitude compared with the maximum concentration (Fig. 5) and were not statistically significant ($P > 0.16$, ANOVA) for the sample size (power < 0.37). However, the interaction between the range and number of concentrations was statistically significant for AUC ($P < 0.05$, ANOVA), reflecting that a greater number of concentrations in the calibration significantly improved AUC for the lowest concentration range (30 mM), but not the higher concentration ranges [Fig. 5(d)]. Increasing the range and number of concentrations in the calibration from the minimum (30 × 2) to maximum (90 × 7) improved the median RMSE, specificity, and AUC from ~2.5 to ~1.5 mM, ~0.75 to ~0.90, and ~0.85 to ~0.93, respectively, while the median sensitivity decreased from ~0.87 to ~0.80 (Fig. 5). Thus, the median RMSE of material decomposition was similar in magnitude to that measured by ICP-OES on the same samples [Fig. 5(a) and Table II].

3.D. Scaling factor calibration

The effects of the calibration method (Table I) on the scaling factor calibration and subsequent material decomposition were further evaluated using a fixed calibration method (e.g., 30 × 2 and 90 × 7) for the material basis matrix (Fig. 6). Quantitative evaluation of material decomposition revealed

TABLE III. MANOVA for the overall effects of the range and the number of gadolinium contrast agent concentrations on the quantitative (RMSE) and spatial (sensitivity, specificity and AUC) accuracy of material decomposition for each possible combination of calibration methods for both the material basis matrix and scaling factor (Table I). Statistically significant effects are shown by italicized *P*-values.

	Material basis matrix calibration			Scaling factor calibration		
Metric	Maximum concentration	Number of concentrations	Interaction	Maximum concentration	Number of concentrations	Interaction
RMSE	*<0.0001*	0.10	*<0.0001*	*<0.0001*	*<0.005*	*<0.0001*
Sensitivity	*<0.0001*	0.79	0.07	*<0.0001*	*<0.0001*	*<0.0001*
Specificity	*<0.0001*	*<0.005*	*<0.0001*	*<0.001*	0.09	*<0.001*
AUC	*<0.0001*	*<0.0001*	*<0.001*	0.68	0.47	0.92

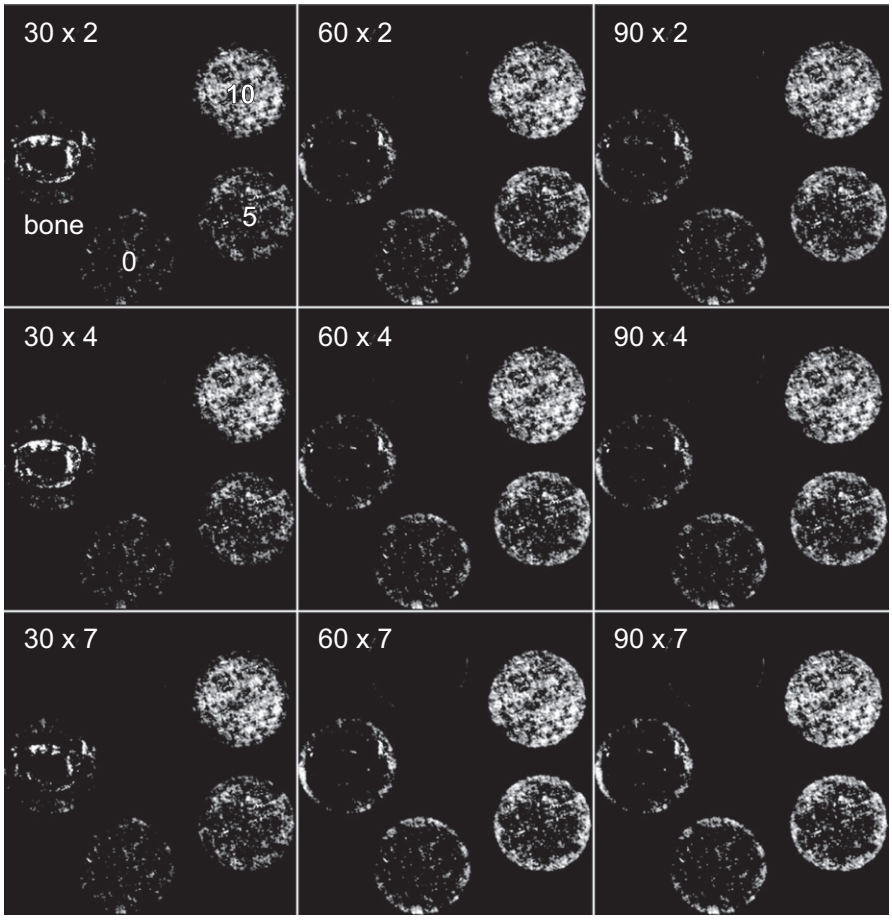

FIG. 4. Representative gadolinium segmented images of the sample phantom after material decomposition using the same calibration method (Table I) for both the material basis matrix and scaling factor. The minimum and maximum threshold for gadolinium segmentation was set at 0 and 20 mM, respectively. Note that for evaluation of accuracy, true positive VOIs included the two concentrations of gadolinium, while true negative VOIs included the water and bone samples.

that the effects of the scaling factor calibration (Fig. 6) were smaller in magnitude compared with the material basis matrix calibration (Fig. 5). Nonetheless, several trends were apparent. Scaling factor calibration using a greater concentration range significantly increased RMSE and sensitivity for lowest concentration range (30 mM), but decreased RMSE and sensitivity for the higher concentration ranges (60 and 90 mM) [$P < 0.05$, ANOVA, except $P = 0.07$ in Fig. 6(a)]. Similarly, an increased number of concentrations in the scaling factor calibration increased RMSE and sensitivity for the lowest concentration range (30 mM), but decreased RMSE and sensitivity for the higher concentration ranges (60 and 90 mM); however, these effects were not statistically significant ($P > 0.37$, ANOVA) for the sample size (power < 0.21), except for sensitivity using the 30 $\times$ 2 basis matrix calibration [$P > 0.05$, ANOVA, Fig. 6(b)].

4. DISCUSSION

4.A. Calibration phantoms

Differences between the measured and expected gadolinium concentration in calibration phantoms were systematic (mean error $\approx$ RMSE $\approx$ 3 mM). The accuracy of ICP-OES calibrated by certified standards is typically on the order of 1–10%.[26] Therefore, the measured differences were most likely due to expected error in ICP-OES combined with random error in the preparation of the gadolinium stock solution from which aliquots were subsequently diluted for each gadolinium concentration. For comparison, the RMSE of the measured x-ray attenuation used in multiple linear regressions (Table II), ranged from 0.4 to 2.7 mM in the three lowest energy bins and 1.3 to 10.7 mM in the two highest energy bins, as estimated by converting the RMSE from HU (Table II) to mM using the slope of the linear regression. Moreover, the RMSE of the measured x-ray attenuation of gadolinium concentrations in the calibration phantoms was typically < 30 HU. A differential contrast (ΔHU) of 30 HU has been suggested to be the threshold for visible perception.[27] A greater RMSE for higher energy bins of the photon-counting detector (Table II) was expected due to decreased photon counts from the x-ray source with increased photon energy. Moreover, a deviation from linearity was evident in the relationship between x-ray attenuation and gadolinium concentration for the 60.1–73.1 and 73.1–120 keV energy bins at concentrations greater than 60 mM (Fig. 3), where low photon counts were exacerbated by higher concentrations absorbing more photons and further reducing photon counts at the detector.

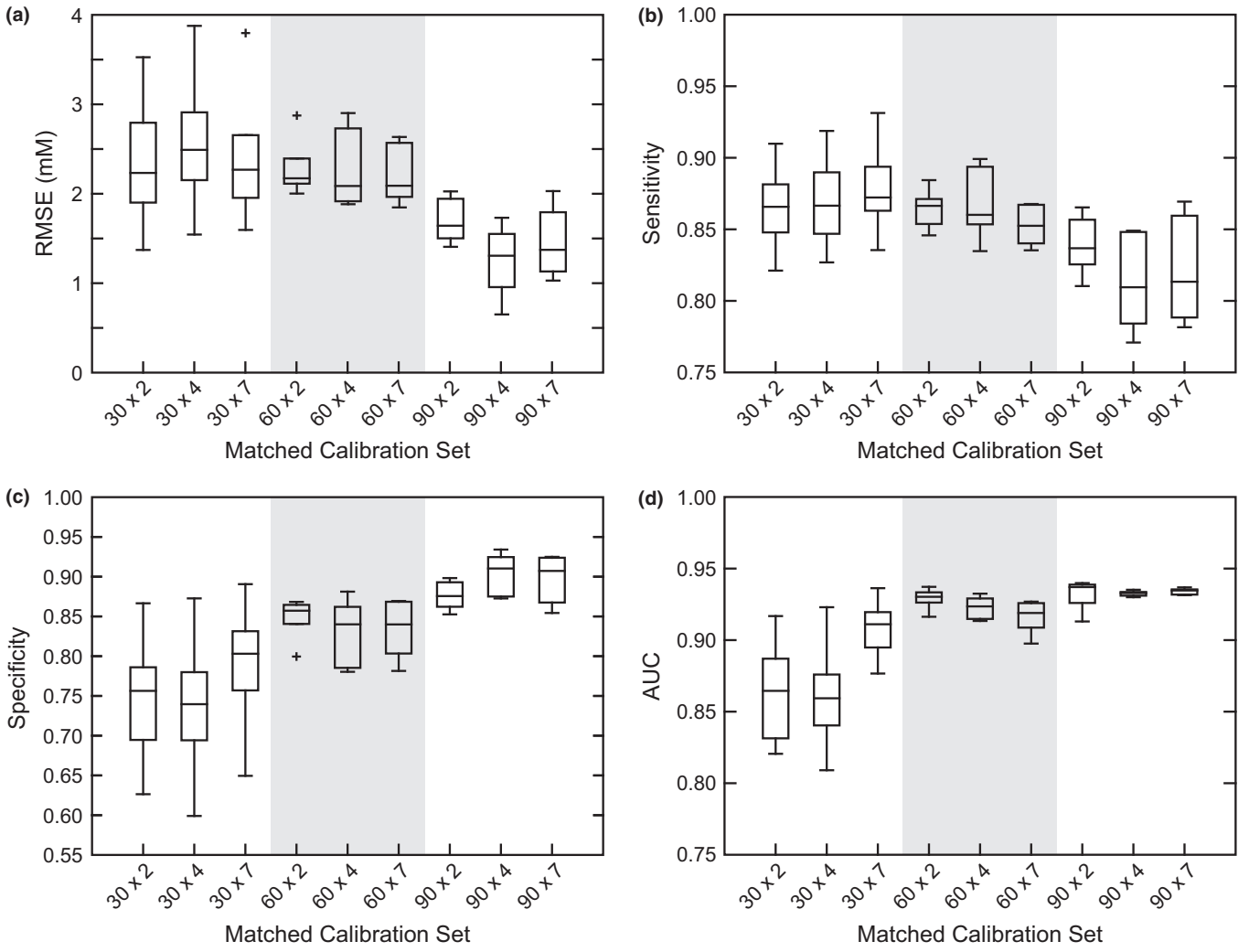

FIG. 5. Boxplots showing the (a) RMSE, (b) sensitivity, (c) specificity, and (d) AUC of the resulting quantitative material decomposition of gadolinium using the same calibration method (Table I) for both the material basis matrix and scaling factor. The box and whiskers show the 10th, 25th, 50th (median), 75th, and 90th percentiles. The overall effects of the range and the number of gadolinium concentrations in the calibration are shown in Table III.

4.B. Quantitative accuracy of material decomposition

Photon-counting spectral CT was able to achieve a median RMSE as low as ~1.5 mM (~0.24 mg/mL) for quantifying a gadolinium contrast agent [Figs. 5(a) and 6(c)], which was similar in magnitude to that measured for gadolinium in previous studies[19,28] and to that measured by ICP-OES on the same samples in this study [Fig. 5(a) and Table II]. Therefore, the quantitative accuracy of material decomposition using photon-counting spectral CT was comparable to that of optical spectroscopic techniques commonly used to measure mM (and lower) elemental concentrations in media. This suggests potential for using photon-counting spectral CT in analytical applications that extend beyond diagnostic imaging in biomedicine. In diagnostic imaging, an RMSE of ~1 to 1.5 mM is desirable to improve the signal-to-noise ratio to ~3 for an observable signal of ~5 mM (~0.8 mg/mL gadolinium) in this study.

4.C. Effects of calibration methods on material decomposition

The results of this study demonstrate important effects of calibration methods on the accuracy of quantitative material decomposition using photon-counting spectral CT. The material basis matrix calibration was more sensitive to changes in the calibration methods than the scaling factor calibration (Table III, Figs. 5 and 6). The material basis matrix calibration significantly influenced both the quantitative (RMSE) and spatial (AUC) accuracy of material decomposition, while the scaling factor calibration influenced RMSE but not AUC (Table III). Note that AUC is a robust measure of the overall spatial accuracy based on the relative gain in specificity with decreasing sensitivity over a range of thresholds.[29] The material basis matrix and scaling factor calibrations might be expected to exhibit a primary influence on the spatial (AUC) and quantitative (RMSE) accuracy of material decomposition, respectively, based on their implementation. However,

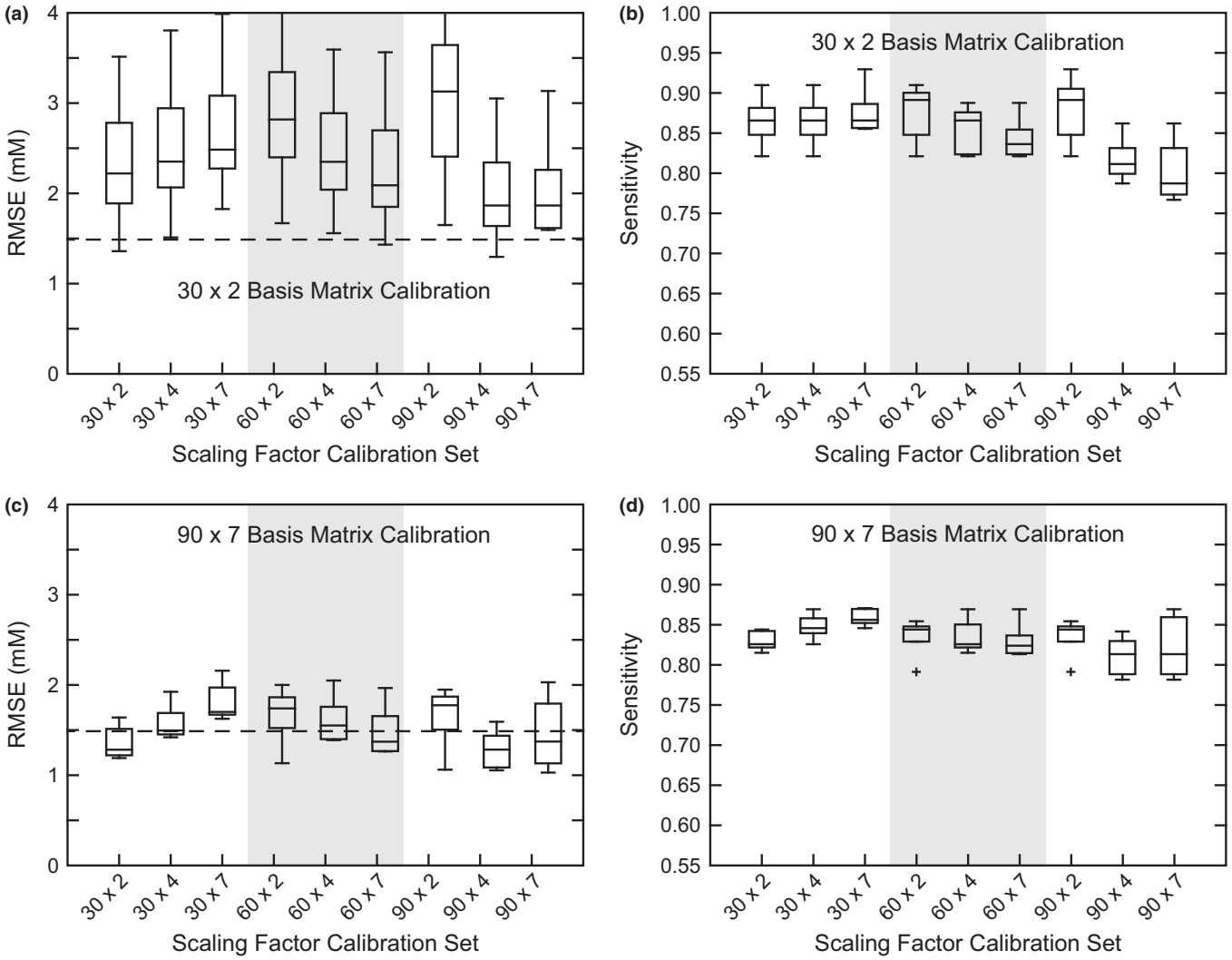

FIG. 6. Boxplots showing the effects of the calibration method (Table I) for the scaling factor calibration on the (a,c) RMSE and (b,d) sensitivity of the resulting quantitative material decomposition of gadolinium using a fixed calibration method for the material basis matrix: (a,b) 30 × 2 and (c,d) 90 × 7. The box and whiskers show the 10th, 25th, 50th (median), 75th, and 90th percentiles. The dashed line shows the RMSE corresponding to detection of 5 mM gadolinium with a signal-to-noise ratio of 3. The overall effects of the range and the number of gadolinium concentrations in the calibration are shown in Table III.

the material basis matrix not only identifies the spatial location of a material of interest in the decomposed image, but also provides an initial estimation of the fractional abundance [Eqs. (1)–(3)], and thus determines the variance in fractional abundance of a material of interest. The scaling factor subsequently converts the fractional abundance to concentration using the linear regression models. Therefore, the material basis matrix calibration exhibits a primary influence on both the quantitative (RMSE) and spatial (AUC) accuracy of material decomposition.

Both the quantitative (RMSE) and spatial accuracy (AUC) of material decomposition were most improved by using an increased maximum concentration (range) of the gadolinium contrast agent in the basis matrix calibration (Fig. 5). A small decrease in sensitivity was observed for an increased concentration range, but this effect was outweighed by greater improvements in RMSE and specificity, which was further confirmed by improved AUC as a measure the overall spatial accuracy (Fig. 5). Therefore, these results suggest that an increased maximum concentration of the gadolinium contrast agent in the calibration phantom provided a greater signal-to-noise ratio during calibration, which subsequently improved the accuracy of quantitative material decomposition.

An increased number of gadolinium contrast agent concentrations in the material basis matrix and scaling factor calibration was able to improve the quantitative (RMSE) and spatial (AUC) accuracy, respectively, of material decomposition (Table III), but the magnitude of these effects was relatively small compared to the effects of the concentration range and not always statistically significant (Figs. 5 and 6). An increased number of concentrations (or samples) was expected to reduce variance in the linear regression models, improving the accuracy of the initial estimation of the fractional abundance from the material basis matrix and the subsequent material decomposition. However, the MAP estimator simultaneously evaluates the linear regression for

each energy bin such that the overall signal-to-noise ratio of the calibration, which was maximized at the highest concentration range, primarily governed the spatial and quantitative accuracy of material decomposition. Therefore, the results of this study suggest that calibration methods for quantitative material decomposition may not require more than one concentration (for a two-point calibration) of a material of interest. On the other hand, the results also suggest that a low spatial accuracy (specificity and AUC) of material decomposition, due to a low (30 mM) concentration range in the calibration, can be mitigated by using an increased number of concentrations [Fig. 5(d)].

Effects of the calibration methods on the scaling factor calibration were also noteworthy, despite their smaller magnitude compared with the material basis matrix. For a material basis matrix calibration that performed relatively poor (30×2), increasing both the range and number of concentrations in the scaling factor calibration improved (decreased) RMSE, but also decreased sensitivity, compared to using the same range and number of concentrations in the scaling factor calibration [Figs. 6(a) and 6(b)]. However, note that increasing only the range or number of concentrations in the scaling factor calibration resulted in increased RMSE and no change in sensitivity. For a material basis matrix calibration that performed relatively well (90×7), decreasing both the range and number of concentrations in the scaling factor calibration decreased variability in the measure RMSE and sensitivity, compared to using the same range and number of concentrations in the scaling factor calibration [Figs. 6(c) and 6(d)].

Finally, overall variations in calibration methods that improved quantitative accuracy (RMSE) compromised sensitivity, and vice versa (Fig. 5). Moreover, RMSE and sensitivity were the only metrics significantly influenced by the scaling factor calibration at a fixed basis matrix calibration (Fig. 6). These results suggest that selection of the most appropriate of calibration methods may be application-specific. For example, in applications where high sensitivity is desired — perhaps the detection of breast microcalcifications, metastatic cancer, or kidney stones — calibration methods should utilize a lower maximum concentration (range) and greater number of concentrations. Thus, to detect an expected concentration of ~5 mM at a signal-to-noise ratio of 3, calibration methods might include five concentrations spanning 0 to ~10 mM. On the other hand, in applications where quantitative precision and accuracy (RMSE) is desired — perhaps bone densitometry or imaging the biodistribution of a targeted contrast agent — calibration methods should utilize a higher maximum concentration and require a fewer number of concentrations.

4.D. Limitations

Several limitations should be considered when interpreting the results of this study and their implications. Material decomposition was performed in the image domain rather than the projection domain. While the implementation of image- and projection-based decomposition is fundamentally different, the effects of calibration methods should be similar for either approach provided that all other factors (e.g., photon counts) are similar. Therefore, the general trends in the results of this study are expected to be broadly applicable to supervised material decomposition in both the image and projection domain.

Evaluation of the quantitative and spatial accuracy of material decomposition included only the decomposition of gadolinium to simplify the interpretation of results. However, the hard-constrained MAP estimator requires a fully defined material basis matrix to produce accurate results. Therefore, fluctuations within the vector of values used to define the gadolinium signal can adversely affect the results of other material channels. All material channels should be considered simultaneously in order to fully evaluate the material basis matrix calibration. Moreover, decomposition of gadolinium and calcium was performed separately. Therefore, future studies should investigate mixed combinations of two materials (e.g., gadolinium and calcium) for both calibration and evaluation.

True positive and true negative sample sizes were kept equal for evaluation of spatial accuracy (sensitivity, specificity, and AUC). However, the same sample sizes were used for evaluation of quantitative accuracy (RMSE) such that the zero concentration samples were more heavily weighted due being equivalent in size to the 5 and 10 mM gadolinium concentrations combined. Therefore, RMSE measurements may be either under- or overestimated depending on whether residuals at zero concentration were lower or higher, respectively, compared with higher concentrations. This effect could be minimized by increasing the number of samples within the sample phantom concentration range.

Beam hardening was not corrected prior to material decomposition, such that artifacts were observable within the lowest energy bin in voxels adjacent to the highly attenuating bone sample. Beam hardening artifacts had minimal effect on the results of this study, but would be expected to have a stronger influence on the material decomposition for a larger specimen size. Thus, a beam hardening correction may need to be implemented for larger specimens, e.g., animal or human subjects

Finally, the magnitude of metrics used to evaluate the spatial and quantitative accuracy of material decomposition is dependent on the concentrations in the sample phantom. In this study, the sample phantom contained 0, 5, and 10 mM gadolinium concentrations that were selected to cover the expected range that would be observable *in vivo*. However, this limited the analysis of sensitivity and specificity to a minimum threshold value of 5 mM. A more robust evaluation of the accuracy of material decomposition closer to the desired detection limit could be obtained by increasing the number of samples analyzed between 0 and 5 mM.

5. CONCLUSIONS

The accuracy of quantitative material decomposition for detecting a gadolinium contrast agent in photon-counting

spectral CT was significantly influenced by calibration methods which must therefore be carefully considered for the intended diagnostic imaging application. The material basis matrix calibration was more sensitive to changes in the calibration methods than the scaling factor calibration. Both the quantitative (RMSE) and spatial (AUC) accuracy of material decomposition were most improved by using an increased maximum gadolinium concentration (range) in the basis matrix calibration. An increased number of gadolinium contrast agent concentrations in the material basis matrix and scaling factor calibration was also able to improve the spatial (AUC) and quantitative (RMSE) accuracy, respectively, of material decomposition, but the magnitude of these effects was relatively small compared to the effects of the concentration range and not always statistically significant. Importantly, the median RMSE of material decomposition was as low as ~1.5 mM (~0.24 mg/mL gadolinium), which was similar in magnitude to that measured by optical spectroscopy on the same samples.

ACKNOWLEDGMENTS

This research was supported by grants from the National Science Foundation (DMR-1309587), Kelly Cares Foundation, St. Joseph Regional Medical Center, and the University of Notre Dame Equipment Renewal and Restoration Program. The authors acknowledge the Notre Dame Center for Environmental Science and Technology (CEST) for ICP-OES and that one aspect of the software used in this study was initiated by T. E. Curtis at the Physiological Research Imaging Laboratory at the Mayo Clinic College of Medicine. The authors thank Ken Sauer at the University of Notre Dame for helpful conversations regarding this work.

[a]Author to whom correspondence should be addressed. Electronic mail: rroeder@nd.edu; Telephone: +1 (574) 631 7003.

REFERENCES

1. Kircher MF, Willmann JK. Molecular body imaging: MR imaging, CT, and US. Part I. Principles. *Radiology.* 2012;263:633–643.
2. Nam SY, Ricles LM, Suggs LJ, Emelianov SY. Imaging strategies for tissue engineering applications. *Tissue Eng Part B Rev.* 2015;21:88–102.
3. Kalender WA. X-ray computed tomography. *Phys Med Biol*. 2006;51: R29–R43.
4. Hubbell JH, Seltzer SM. *Table of X-ray Mass Attenuation Coefficients and Mass Energy-Absorption Coefficients (Version 1.4)*. Gaithersburg, MD: National Institute of Standards and Technology; 2004. Available at http://physics.nist.gov/xaamdi.
5. Hubbell JH, Gimm HA, Overbo I. Pair, triplet, and total atomic cross sections (and mass attenuation coefficients) for 1 MeV–100 GeV photons in elements $Z = 1$ to 100. *J Phys Chem Ref Data*. 1980;9:1023–1148.
6. He P, Yu H, Bennett J, et al. Energy-discriminative performance of a spectral micro-CT system. *J X-Ray Sci Technol*. 2013;21:335–345.
7. Graser A, Johnson TR, Chandarana H, Macari M. Dual energy CT: preliminary observations and potential clinical applications in the abdomen. *Eur Radiol*. 2009;19:13–23.
8. Johnson TRC, Kraub B, Sedlmair M, et al. Material differentiation by dual energy CT: initial experience. *Eur Radiol*. 2007;17:1510–1517.
9. Yu Z, Leng S, Jorgensen S, et al. Initial results from a prototype whole-body photon-counting computed tomography system. *Proc SPIE Int Soc Opt Eng*. 2015;9412:94120W.
10. Roessl E, Proksa R. K-edge imaging in X-ray computed tomography using multi-bin photon counting detectors. *Phys Med Biol*. 2007;52:4679–4696.
11. Yu H, Xu Q, He P, et al. Medipix-based spectral micro-CT. *CT Li Lun Yu Ying Yong Yan Jiu*. 2012;21:583.
12. Jorgensen SM, Vercnocke AJ, Rundle DS, Butler PH, McCollough CH, Ritman EL. Evaluation of a photon counting Medipix 3RX CZT spectral X-ray detector. *Proc SPIE Int Soc Opt Eng*. 2016;9969:99690J.
13. Alvarez RE, Macovski A. Energy-selective reconstructions in X-ray computerised tomography. *Phys Med Biol*. 1976;21:733–744.
14. Alessio AM, MacDonald LR. Quantitative material characterization from multi-energy photon counting CT. *Med Phys*. 2013;40:031108.
15. Le HQ, Molloi S. Least squares parameter estimation methods for material decomposition with energy discriminating detectors. *Med Phys*. 2011;38:245–255.
16. Barber RF, Sidky EY, Schmidt TG, Pan XC. An algorithm for constrained one-step inversion of spectral CT data. *Phys Med Biol*. 2016;61:3784–3818.
17. Weidinger T, Buzug TM, Flohr T, Kappler S, Stierstorfer K. Polychromatic iterative statistical material image reconstruction for photon-counting computed tomography. *Int J Biomed Imaging*. 2016;2016:5871604.
18. Schirra CO, Roessl E, Koehler T, et al. Statistical reconstruction of material decomposed data in spectral CT. *IEEE Trans Med Imaging*. 2013;32:1249–1257.
19. Clark DP, Badea CT. Spectral diffusion: an algorithm for robust material decomposition of spectral CT data. *Phys Med Biol*. 2014;59:6445–6466.
20. Alvarez RE. Efficient, non-iterative estimator for imaging contrast agents with spectral X-ray detectors. *IEEE Trans Med Imaging*. 2016;35:1138–1146.
21. Nakada K, Taguchi K, Fung GSK, Amaya K. Joint estimation of tissue types and linear attenuation coefficients for photon counting CT. *Med Phys*. 2015;42:5329–5341.
22. Deuerling JM, Rudy DJ, Niebur GL, Roeder RK. Improved accuracy of cortical bone mineralization measured by polychromatic microcomputed tomography using a novel high mineral density composite calibration phantom. *Med Phys*. 2010;37:5138–5145.
23. Cole LE, Ross RD, Tilley JMR, Vargo-Gogola T, Roeder RK. Gold nanoparticles as contrast agents in X-ray imaging and computed tomography. *Nanomedicine*. 2015;10:321–341.
24. Berger MJ, Hubbell JH, Seltzer SM, et al. *XCOM: Photon Cross Section Database (Version 1.5)*. Gaithersburg, MD: National Institute of Standards and Technology; 2010. Available at http://physics.nist.gov/xcom.
25. Themelis K, , Rontogiannis AA. A soft constrained MAP estimator for supervised hyperspectral signal unmixing. In: *16th European Signal Processing Conference (EUSIPCO)*. IEEE: Lausanne, Switzerland; 2008.
26. Olesik JW. Inductively coupled plasma-optical emission spectroscopy. In: Brundle CR, Evans CA Jr, Wilson S, eds. *Encyclopedia of Materials Characterization: Surfaces, Interfaces, Thin Films*. Boston: Butterworth-Heinemann; 1992:633–644.
27. Krause W. Delivery of diagnostic agents in computed tomography. *Adv Drug Deliv Rev.* 1999;37:159–173.
28. van Hamersvelt RW, Willemink MJ, de Jong PA, et al. Feasibility and accuracy of dual-layer spectral detector computed tomography for quantification of gadolinium: a phantom study. *Eur Radiol*. 2017, in press. https://doi.org/10.1007/s00330-017-4737-8
29. Rangayyan RM. *Biomedical Image Analysis*. Boca Raton: CRC Press; 2005:1135–1138.