Towards local progression estimation of pulmonary emphysema using CT

M. Staring[a] and M. E. Bakker
Department of Radiology, Division of Image Processing, Leiden University Medical Center, PO Box 9600, 2300 RC Leiden, The Netherlands

J. Stolk
Department of Pulmonology, Leiden University Medical Center, PO Box 9600, 2300 RC Leiden, The Netherlands

D. P. Shamonin, J. H. C. Reiber, and B. C. Stoel
Department of Radiology, Division of Image Processing, Leiden University Medical Center, PO Box 9600, 2300 RC Leiden, The Netherlands

(Received 16 December 2011; revised 21 October 2013; accepted for publication 21 October 2013; published 16 January 2014)

Purpose: Whole lung densitometry on chest CT images is an accepted method for measuring tissue destruction in patients with pulmonary emphysema in clinical trials. Progression measurement is required for evaluation of change in health condition and the effect of drug treatment. Information about the location of emphysema progression within the lung may be important for the correct interpretation of drug efficacy, or for determining a treatment plan. The purpose of this study is therefore to develop and validate methods that enable the local measurement of lung density changes, which requires proper modeling of the effect of respiration on density.
Methods: Four methods, all based on registration of baseline and follow-up chest CT scans, are compared. The first naïve method subtracts registered images. The second employs the so-called dry sponge model, where volume correction is performed using the determinant of the Jacobian of the transformation. The third and the fourth introduce a novel adaptation of the dry sponge model that circumvents its constant-mass assumption, which is shown to be invalid. The latter two methods require a third CT scan at a different inspiration level to estimate the patient-specific density-volume slope, where one method employs a global and the other a local slope. The methods were validated on CT scans of a phantom mimicking the lung, where mass and volume could be controlled. In addition, validation was performed on data of 21 patients with pulmonary emphysema.
Results: The image registration method was optimized leaving a registration error below half the slice increment (median 1.0 mm). The phantom study showed that the locally adapted slope model most accurately measured local progression. The systematic error in estimating progression, as measured on the phantom data, was below 2 gr/l for a 70 ml (6%) volume difference, and 5 gr/l for a 210 ml (19%) difference, if volume correction was applied. On the patient data an underlying linearity assumption relating lung volume change with density change was shown to hold (fit $R^2 = 0.94$), and globalized versions of the local models are consistent with global results (R^2 of 0.865 and 0.882 for the two adapted slope models, respectively).
Conclusions: In conclusion, image matching and subsequent analysis of differences according to the proposed lung models (i) has good local registration accuracy on patient data, (ii) effectively eliminates a dependency on inspiration level at acquisition time, (iii) accurately predicts progression in phantom data, and (iv) is reasonably consistent with global results in patient data. It is therefore a potential future tool for assessing local emphysema progression in drug evaluation trials and in clinical practice. © *2014 American Association of Physicists in Medicine*. [http://dx.doi.org/10.1118/1.4851535]

Key words: emphysema, image registration, local disease progression estimation, dry sponge model

1. INTRODUCTION

Lung densitometry on chest CT images can be used as a surrogate marker for measuring the destruction of lung tissue in pulmonary emphysema.[1] While methods to perform global quantification of lung density exist,[1] local quantification is studied less frequently. Information about the location of emphysema progression may help in assessing the efficacy of drug treatment.[2] For example, with global density quantification a treatment effect could go unnoticed if a drug protects the healthy part of the lungs, while the affected part cannot be saved. In addition, pulmonary emphysema is sometimes confined to one lobe or part of the lobe, which need to be identified for deciding on a treatment plan. Therefore, precise localization of density changes is required. Since emphysema is a slowly progressive disorder, visual inspection of CT scans can only indicate progression in severe cases or after many years, thus a quantitative approach is needed.

Examples of global quantification methods are lung density estimators, which include the average density of the lung, the pth percentile point (Perc, the Hounsfield Unit (HU) at which p percent of the lung voxels have a lower value, e.g., Perc15) or relative area (RA, the percentage of lung voxels below a certain HU, e.g., RA-910). Comparison of these estimators between baseline and follow-up can be used to evaluate disease progression[3] for the complete lung. In the early development of local measurements, these global estimates were used to measure progression in separate partitions stacked cranio-caudally.[2] From these partitions, a rough estimation was given on the predominant location of emphysema (basal or apical). Another, more anatomy-based strategy to locally estimate changes in density is to segment regions of the lung, such as lung lobes[4–7] or even lobe segments.[8] The density estimators can subsequently be computed per region. To obtain a progressively more localized analysis, increasingly smaller objects have to be identified, and the corresponding region in the follow-up scan has to be found. In the end, the density values of voxels have to be compared directly without computing derived parameters like RA and Perc15.

Other algorithms compute CT derived local features based on texture.[9] Pairs of inhale-exhale scans have also been used to measure different components of COPD, such as trapped air: employing thresholds,[10] using RA,[11] the HU ratio between inhale and exhale at corresponding locations,[12] or many more features derived from these pairs in combination with a classifier.[13] Arzhaeva *et al.*[14] proposed a method for estimating progression of interstitial lung disease, using among others a registration-derived subtraction image. Change was marked in three categories as stable, progressed, or regressed. Commonly these methods do not take the inspiration level of the lung into account (no volume correction), so they do not lead to a more accurate local progression estimate of the emphysema component; they have the lobe as the smallest definition of regional progression,[11–13] or only evaluate their methods against global progression measures, pulmonary function tests, or in coarse categories (e.g., GOLD stages).

In this paper, we rely on image registration to establish local correspondence between follow-up chest CT scans. In principle, this enables computation of disease progression on a per-voxel basis. There has been considerable interest in the registration of CT scans of the lungs, either for follow-up or over the respiratory cycle. Gorbunova *et al.*[15] and Yin *et al.*[16] proposed to use a similarity metric based on the sum of squared differences, modified to account for differences in inspiration level. Features other than image intensity have been proposed to improve the registration, such as using landmarks,[17,18] or vessel centerlines and lung surfaces.[19] In 2010, a grand challenge on pulmonary image registration was organized, the EMPIRE10 challenge,[20] which provides a good overview of the many registration approaches. Intensity-based registration algorithms were among the ones with the best registration accuracy.

The major challenge for proper estimation of emphysema progression using CT is the large influence of the inspiration level on lung density, relative to the influence of emphysema. Scanning with the same volume is difficult: even with spirometric control[21] lung volumes can differ 30%, which indicatively translates to 20 gr/l, which is about ten times the expected global progression in emphysema per year[3] (averaged over the entire lungs and over a large population). Other challenges are the inflow of blood during inspiration, which changes lung mass and density as well.[22]

In summary, local progression estimation is confronted with two major challenges: (a) to obtain a sufficient quality of lung tissue correspondence between baseline and follow-up CT scans and (b) to develop a model that best represents the *in vivo* relation between lung density and volume during the respiratory cycle. It is then the aim of this study to develop methods for detailed local analysis of emphysema progression that are independent of the respiratory state of the lungs at acquisition. The methods are based on image registration, are targeted to quantification instead of classification, and provide a measure of lung tissue destruction. This paper is a major extension of our earlier work.[23] We propose a first method based on the assumption that the lung behaves as a (dry) sponge. This method is subsequently modified to allow for a more flexible relation between lung volume and density, based on a third CT scan, resulting in two more methods. Combined with a naïve model that neglects the dependency on lung volume this results in four methods, which are described in Sec. 2. The experiments and results are described in Sec. 3, and a discussion is given in Sec. 4.

2. METHODS

Our goal is to estimate emphysema progression between two time points, $t \in \{b, f\}$, denoting baseline and follow-up, given one or more CT scans for each time point t. The CT scans are denoted by $I_t(\boldsymbol{x})$, where $\boldsymbol{x}$ denotes spatial position within the image. In this image, density is estimated in gr/l by adding 1000 to the Hounsfield Units,[24] as used in clinical trials.[1,25] The basic relation between mass m and volume V is given by $m = V\rho$, where ρ denotes density, which can be rewritten to

$$\log \rho = \log m - \log V, \tag{1}$$

see Fig. 1 for a visual representation. The volumes can be globally measured using lung segmentation, as has been done for global lung densitometry. Local volumes and more importantly the change in local lung volume can also be determined, using the method described in Sec. 2.B. Given this lung volume (change) and the density as measured by the CT scanner, lung mass can be estimated (locally) using a model like Eq. (1) relating mass, volume, and density.

In the remainder of this section, we describe in Sec. 2.A a preprocessing step for data recalibration. Section 2.B describes the image registration technique that finds correspondence of the CT scans between time points. There we also describe how local volume change is commonly derived from the registration result. In Sec. 2.C, we propose four lung models for progression estimation, and we give the main equation for local emphysema progression estimation. Section 2.D describes the visualization of the local progression images used in the paper.

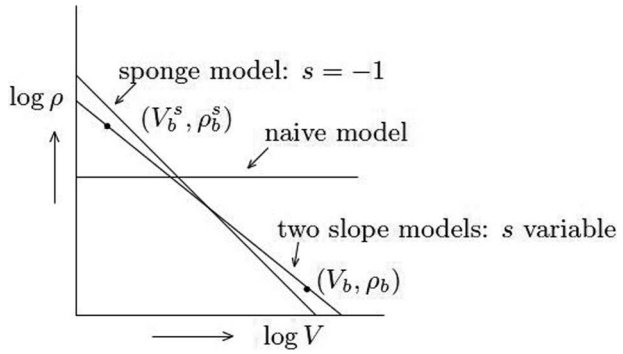

FIG. 1. The relation between density and volume in a log-log plot depicted at baseline. The slope s of the log-log plot is possibly different from -1 and dependent on the location.

2.A. CT image recalibration

Over time the image acquisition characteristics of a CT scanner may change,[26] despite the inscanner calibration for water and air, and therefore the baseline and follow-up scan may not be directly comparable. Since the goal of this work is to compare scans from different time points, it is very important to recalibrate the data. Recalibration is based on automatic measurement of the average HU of air outside the body and of blood in the descending aorta, using the inhouse developed software called Pulmo (commercialized by Medis specials, Leiden, The Netherlands). These measurements should be equal to -1000 and 50 HU, respectively.[26] The intensity values of the scans are rescaled linearly according to these measurements.

2.B. Image registration

To be able to compare corresponding locations of two density images, registration of the original CT data is used. The spatial transformations are estimated using nonrigid intensity-based image registration, using the software package `elastix`,[27] see http://elastix.isi.uu.nl. Image registration is formulated as an optimization problem, where the optimal transformation parameters $\widehat{\boldsymbol{\mu}}$ are found by solving

$$\widehat{\boldsymbol{\mu}} = \arg\min_{\boldsymbol{\mu}} \mathcal{C}(\boldsymbol{T}_{\boldsymbol{\mu}}; I_b, I_f), \tag{2}$$

with $\boldsymbol{T}_{\boldsymbol{\mu}}$ the coordinate transformation parameterized by $\boldsymbol{\mu}$, which relates the two images. To obtain a course alignment, an affine registration is performed prior to nonrigid registration. The nonrigid transformation is modeled by B-splines.[28] Several similarity measures were tested, i.e., sum of squared differences (SSD), normalized correlation (NC), mutual information (MI) and its normalized version (NMI). In addition, localized versions[29] of these metrics were tested (LSSD, LNC, LMI, LNMI), by drawing samples from a randomly chosen cubic sample region. A region size of $50 \times 50 \times 50$ mm was used in this paper.[29] A multiresolution approach was used for images as well as transformations, with a Gaussian image pyramid. An adaptive stochastic gradient descent (ASGD) optimizer[30] was used for solving Eq. (2). Lung masks obtained from the software Pulmo were used to focus the registration on the lungs. Lung masks were generated by means of a standard region growing algorithm, described by Stoel and Stolk,[21] and excludes the large vessels. Exact registration parameter settings can be found on the parameter file database hosted at the `elastix` website (http://elastix.isi.uu.nl , entry `par0015`). For all registrations, we used 1000 iterations per resolution, and if not stated otherwise five resolutions and a final B-spline grid spacing of 10 mm were employed.

The determinant of the spatial Jacobian of the transformation $J_{\boldsymbol{T}} = \det(\partial \boldsymbol{T}/\partial \boldsymbol{x})$ is a measure of the local relative change in volume between baseline and follow-up (V_f/V_b, where V_b and V_f are the local volumes at baseline and follow-up, respectively). A value equal to 1 means no change in volume; >1 means expansion; <1 means compression; <0 indicates a folding in the deformation field. The measure is quantitative: a value of 1.1 means a 10% increase in volume. It has been shown that the spatial Jacobian correlates with ventilation[31] and pulmonary function.[32] In Sec. 2.C, we describe how it is used in our methods.

2.C. Lung models and progression measure

We developed a number of models that describe the relation between density and volume in the lung, and calculate progression from them.

2.C.1. Uncorrected model

A naïve model would be to assume that density is independent of volume changes and to measure progression by a simple subtraction of the matched images, implying the lung model

$$I_f(\boldsymbol{T}(\boldsymbol{x})) = I_b(\boldsymbol{x}) + \text{progression}(\boldsymbol{x}). \tag{3}$$

2.C.2. Sponge model

The (dry) sponge model of the lung assumes mass preservation over the lung breathing cycle $m_b = m_f$, so $\rho_f = \rho_b V_b / V_f$, or

$$I_f(\boldsymbol{T}(\boldsymbol{x})) = I_b(\boldsymbol{x})[J_{\boldsymbol{T}}(\boldsymbol{x})]^{-1} + \text{progression}(\boldsymbol{x}). \tag{4}$$

This method was originally applied in 2003, with a volume correction term globally estimated from segmented lung volumes.[34]

2.C.3. Adapted slope models

The assumption of mass preservation by the sponge model implies a linear relation between $\log V$ and $\log \rho$, with a slope of -1. It is known, however, that this assumption is not correct, due to variability in lung blood perfusion during breathing,[22] trapped air, and CT scanner effects. This changes the measured mass of the lungs. Figure 2 shows the variation in volume-density slope as measured in a patient population. These experimental results[33] suggest to adapt the sponge model, such that the slope s is not fixed to -1. The model is adapted to: $\rho_f = \rho_b (V_f/V_b)^s$, which can be expressed in terms of the CT images by

$$I_f(\boldsymbol{T}(\boldsymbol{x})) = I_b(\boldsymbol{x})[J_{\boldsymbol{T}}(\boldsymbol{x})]^{s(\boldsymbol{x})} + \text{progression}(\boldsymbol{x}), \tag{5}$$

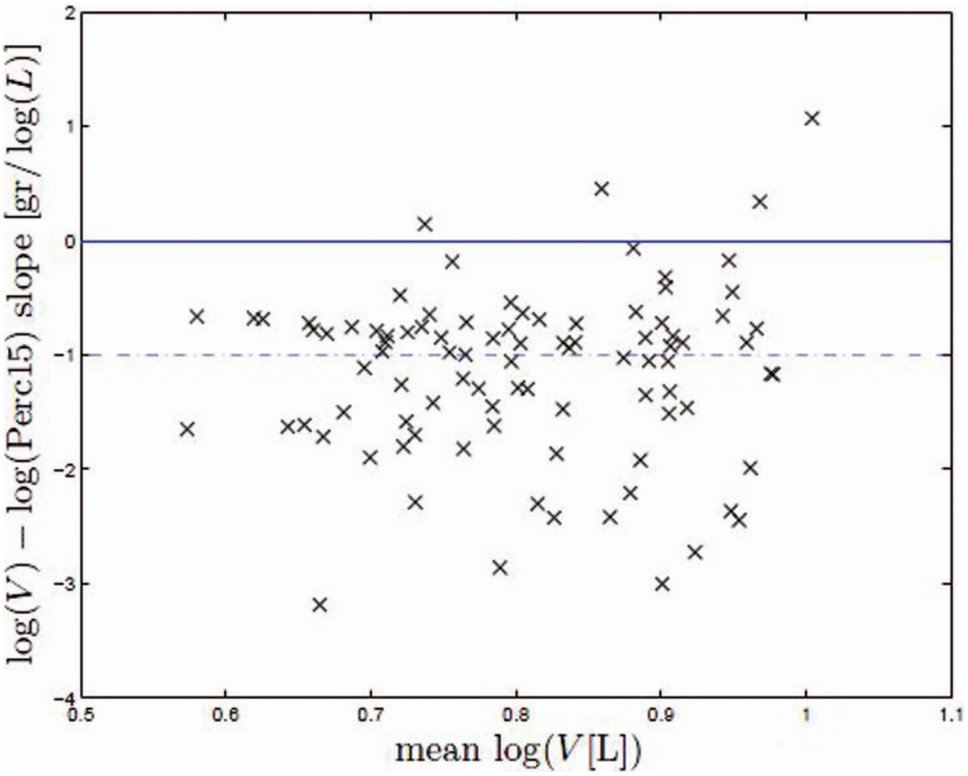

FIG. 2. Volume-density slopes in patient data as measured in a multicenter trial. Vertically the slope of the log of volume against the log of Perc15 is given. Reproduced from Ref. 33.

where $s(\boldsymbol{x})$ is the slope of the log volume-log density relation that may or may not be dependent on the location $\boldsymbol{x}$ (see below). A slope of $s = -1$ corresponds to the sponge model. For $s > -1$, the adapted model states that when inhaling ($V_f > V_b$), the density decreases with a rate below the sponge model, and vice versa. In other words, when inhaling, for $s > -1$, the density decreases less than can be contributed to the increase in volume, so mass is entering the lungs. As in the sponge model, the adapted slope models assume a linear relationship between $\log V$ and $\log \rho$ over the breathing cycle. This assumption is verified in the experiments.

The slope parameter s enables a more realistic model, but requires setting to a proper value. Estimation of this value comes at the cost of an additional scan taken at baseline or follow-up at a different inspiration level, see Fig. 3. We can assume that there is no change in emphysema at the same time point, so we can estimate the (local) relation between density and volume for a given patient. This extra scan is denoted by I_t^{FRC}, with corresponding volume V_t^{FRC} and density ρ_t^{FRC}, where ∼FRC means approximate functional residual capacity.

2.C.3.a. Global slope model. A global but patient-specific value of the slope s (independent of spatial location $\boldsymbol{x}$) can be determined by measuring the total lung volumes and overall mean densities of the scans I_t and I_t^{FRC} and using $s(\boldsymbol{x}) = s = (\log \rho_t^{\mathrm{FRC}} - \log \rho_t)/(\log V_t^{\mathrm{FRC}} - \log V_t)$, see Fig. 1.

2.C.3.b. Local slope model. The slope is estimated locally from the CT scans at baseline time b at position $\boldsymbol{x}$ as follows:

$$s(\boldsymbol{x}) = s_b(\boldsymbol{x}) = \frac{\log I_b^{\mathrm{FRC}}(\boldsymbol{T}_b(\boldsymbol{x})) - \log I_b(\boldsymbol{x})}{\log J_{\boldsymbol{T}_b}(\boldsymbol{x})}. \tag{6}$$

This allows for a local evaluation of the volume-density slope to compensate for local differences in lung perfusion, as perfusion is nonstationary.

In summary, with the above models, local progression in terms of density changes can be estimated by

$$\text{progression}(\boldsymbol{x}) = I_f(\boldsymbol{T}(\boldsymbol{x})) - I_b(\boldsymbol{x})[J_{\boldsymbol{T}}(\boldsymbol{x})]^{s(\boldsymbol{x})}, \tag{7}$$

where $s(\boldsymbol{x}) = 0$ for the uncorrected model, $s(\boldsymbol{x}) = -1$ for the sponge model, $s(\boldsymbol{x}) = s$ (a global but patient-specific constant) for the global slope model, and $s(\boldsymbol{x})$ as estimated with Eq. (6) for the local slope model. A "progression measure" equal to zero indicates no progression, and >0 (<0) indicates more (less) tissue mass.

In order to enable comparison between our local progression estimation and previous global estimates, we also globalized the local measurements by taking the median value over all local data. The median is a robust estimator of the mean local progression, less sensitive to outliers.

2.D. Data visualization

In the remainder of this paper, we visualize the progression image (7) either by showing it as a grayscale image (e.g., Fig. 10), or as a color overlay on the original baseline image (e.g., Fig. 11). The color overlay is constructed as follows: vasculature is detected by a vessel enhancement filter[35]

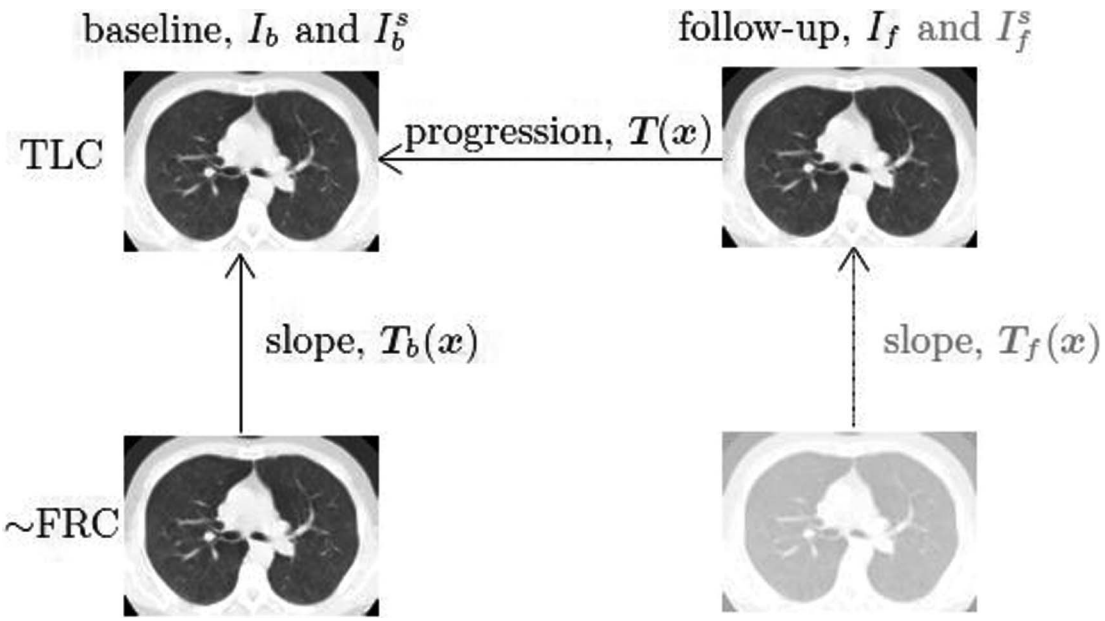

FIG. 3. Relations between chest CT scans. At baseline we have two CT scans available at two different inspiration levels: I_b and I_b^s. Scans are related by the transformations $\boldsymbol{T}(\boldsymbol{x})$, $\boldsymbol{T}_b(\boldsymbol{x})$, and $\boldsymbol{T}_f(\boldsymbol{x})$. TLC: total lung capacity; ~FRC: approximate functional residual capacity.

and depicted in red; a decrease in density (progression of emphysema) is depicted in green; an increase in density in yellow (green and yellow were chosen to account for color-blindness of the user).

3. EXPERIMENTS AND RESULTS

In this section, we present the evaluation of the two main components that enable local progression estimation, i.e., (a) the ability to find correspondence, and (b) the accuracy of the proposed lung models. We optimized the registration procedure on patient data (see Sec. 3.B). The lung models were evaluated on both phantom and patient data, in several steps. First, we performed a phantom experiment to explore the contributions of the intrinsic sources of error, such as interpolation and scanner variability, see Sec. 3.C. We proceed with the evaluation of the lung models, on phantom data in Sec. 3.D, and patient data in Sec. 3.E. The data are described in Sec. 3.A.

3.A. Data

A phantom mimicking the lung was constructed similar to Stoel *et al.*,[36] see Fig. 4. It consists of a cylinder filled with foam representing lung tissue. Volume is adjustable by moving a piston inside the cylinder, while mass is retained. In addition, local density can be modified by segmenting a piece of foam, and digitally modifying the density at that location. The phantom was scanned at several volumes on a Toshiba Aquilion 64 scanner. The CT scan was reconstructed at a 512 × 512 matrix with an inplane resolution of 0.4 × 0.4 mm, using a soft FC04 kernel. Each scan contained ~700 slices, which were 0.5 mm thick, with 0.3 mm increment.

Additionally, we gathered follow-up CT data sets from 21 patients (ten males; aged 34–74 at baseline, mean 56) suffering from pulmonary emphysema, due to $\alpha(1)$-antitrypsin deficiency.[1] These scans were acquired during the SPREAD study[1] with institutional review board approval, using a Toshiba Aquilion 4 scanner with scan parameters: 135 kVp; 20 mAs per rotation; rotation time 0.5 s; collimation: 4 × 5 mm. Scans were made during breath hold, without

FIG. 4. The phantom mimicking the lung.

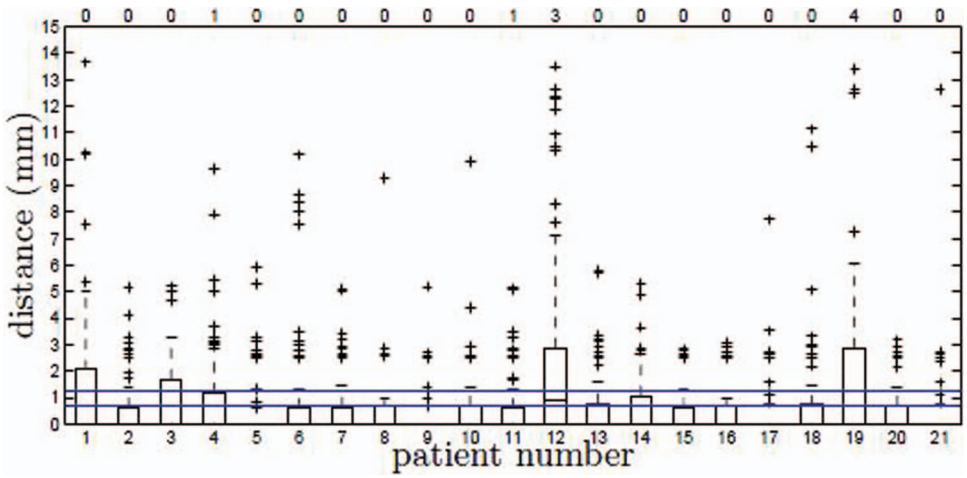

FIG. 5. The interobserver difference on patient data. Numbers on top of the graph denote the number of differences larger than 15 mm. The two horizontal lines denote inplane voxel size (0.7 mm) and half the slice increment (1.25 mm), respectively. Patients 12 and 19 are excluded from further analysis.

contrast media. Images were reconstructed with a standardized protocol optimized for lung densitometry,[21] including a soft FC12 kernel, using a slice thickness of 5 mm and an increment of 2.5 mm, with an inplane resolution of around 0.7×0.7 mm. Note that the thick slice reconstruction was chosen as it followed from an optimization study for global densitometry,[21] and is also used in several drug evaluation trials.[1,25] At each time point (23–34 months apart, median 29), two scans were acquired at different inspiration levels (patients were verbally instructed): total lung capacity (TLC) and ∼FRC.

3.B. Image registration optimization: Patient data

We optimized the registration procedure by determining which registration setup performs the best for intrapatient registration of our data. We tested the effect on registration accuracy of (i) using lung masks to focus the registration; (ii) the similarity metric (SSD, NC, MI, NMI); and (iii) using locality in the metric[29] (LSSD, LNC, LMI, LNMI). The TLC follow-up scan was registered nonrigidly to the TLC at baseline. For initialization, an affine registration using NC was performed prior to nonrigid registration. A Wilcoxon signed rank test was performed to compare the registration results, where $p < 0.05$ denotes a statistically significant difference. The mean error per patient was computed, which were then compared between the different registration setups. Bonferroni adjustment of p was used as a conservative multiple-comparison correction.

The ground truth for the registration was defined using the method by Murphy *et al.*[37] The algorithm automatically finds 100 evenly distributed points in the baseline, only at characteristic locations. Subsequently, corresponding points in the follow-up scan are predicted by the algorithm and shown in a graphical user interface for inspection and possibly correction. Two observers participated in this study, a pulmonologist (J.S.) and an experienced CT researcher (M.E.B.). They first annotated the points independently, after which a consensus reading was performed on those points the observers did not agree exactly. As a result, consensus reading was performed on a median of 29 points per patient (range: 19–39). Points for which an observer could not find a corresponding point were marked as unsure. If one of the observers marked a point as unsure it was considered an unreliable point, and those points were excluded from validation to obtain the ground truth as reliable as possible. All automatically predicted points were manually checked. Two patients (12 and 19) progressed considerably, such that it was not possible to establish correspondence at certain locations (a clear misregistration), and the corresponding CT scans were therefore excluded from further evaluation (except to generate Fig. 8). From the remaining 19 scan pairs we excluded 51 out of a total of 1900 points, of which 31 were from one patient (patient 13). The interobserver difference is depicted in Fig. 5 with a mean and median distance of 0.63 and 0.0 mm, respectively. This means that for the majority of points the observers agreed exactly, which is facilitated by the software prediction strategy to improve the reproducibility, the thick slices that limit the number of landmark choices, and the characteristic locations of the points. Optimization was subsequently performed using the registration error derived from the manually annotated corresponding points.

Figure 6 shows the results for the registration optimization. The use of lung masks slightly but significantly improved the registration for SSD and NC, where median errors decrease from 1.1 to 1.0 mm for both methods, and slightly decreased the accuracy for MI (from 1.1 to 1.4 mm) and NMI (from 1.1 to 1.3 mm), see Fig. 6(a). Localization of the similarity measures significantly worsened the registration results for our application to 6.8, 2.0, 2.0, and 1.9 mm, respectively, notably for the SSD metric [Fig. 6(b)]. SSD and NC with the lung mask were the best measures as indicated by their small median error (both 1.0 mm), and we select NC for further experiments. The median registration error was therefore below half the slice increment for the best registration methods.

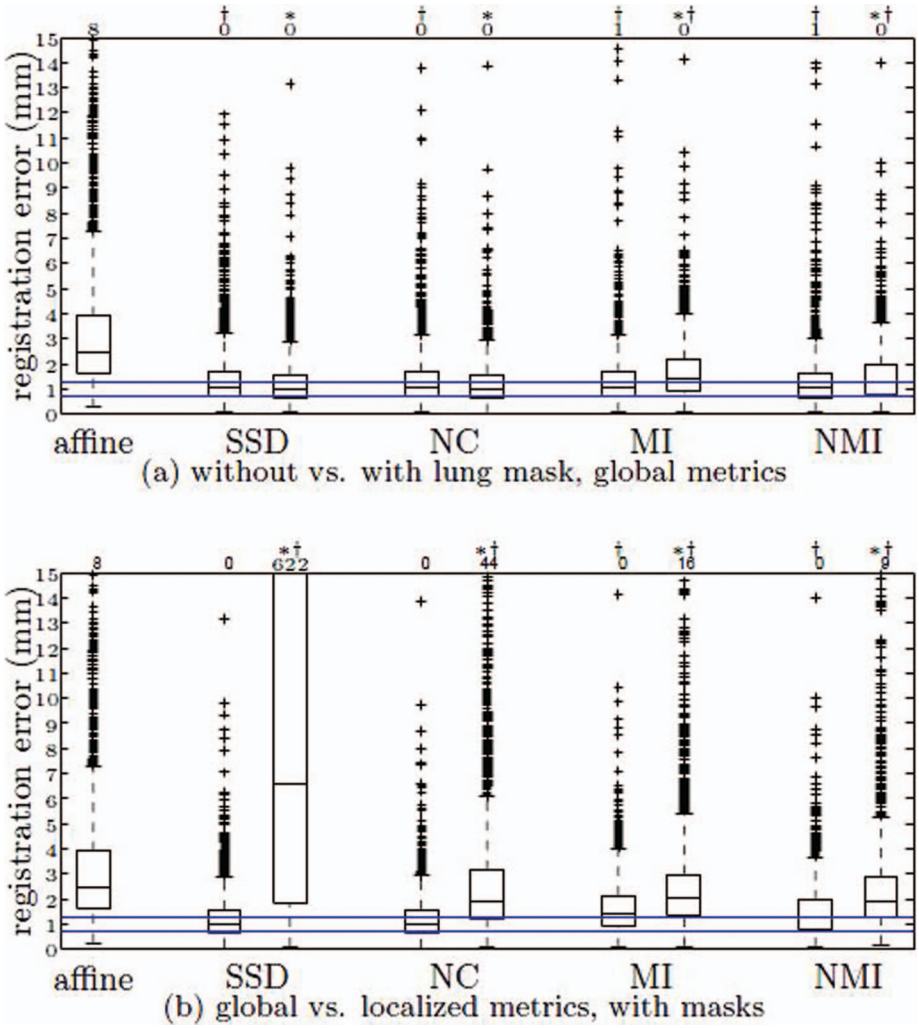

FIG. 6. Registration accuracy for the several methods, evaluated on patient data using 1849 locations. Numbers on top of the graph denote the number of registration errors larger than 15 mm. A star denotes a significant difference with the previous column after Bonferroni correction; a † with NC using a mask. The two horizontal lines denote inplane voxel size (0.7 mm) and half the slice increment (1.25 mm), respectively. For (a), in each group of two boxplots the left one denotes registration without using a lung mask and the right one with. Similarly, for (b), where global vs local metrics are plotted. The affine registration accuracy on the left-most column was added as a benchmark: (a) without vs with lung mask, global metrics, (b) global vs localized metrics, with masks.

3.C. Intrinsic sources of error: Phantom study

To establish the influence of related factors on the progression estimation accuracy, like interpolation, registration, and CT scanner effects, we performed some basic tests on the phantom that do not change volume and density, and ideally should result in measuring no progression: (C1) to evaluate interpolation errors we scanned the phantom, rotated the resulting images digitally over 6°, 29°, and 90° around the vertical axis, and rotated them digitally back; (C2) for estimating errors in the registration process we registered the phantom with itself after an initial digital offset; (C3) to study the combined errors from registration and interpolation, we scanned the phantom, digitally rotated the image, and registered the original image with the rotated version; (C4) to estimate errors from the CT scanner in addition to interpolation and registration errors we scanned the phantom, moved it physically in the scanner, rescanned it, and registered the two images. In the above experiments, a rigid registration was performed using four resolution levels and NC as a similarity measure.

The results are accumulated using the mean and standard deviation of the progression within the phantom, and are given in Table I, for the uncorrected model only, since no change in volume or density was made. Except in experiment C4, small systematic and random errors were found in the progression estimation. So, digital resampling (interpolation) has not much influence on the results and simple registration did not produce large systematic as well as random errors. However, more complex registrations on clinical data may produce larger errors. Experiment C4 showed that the highest variability originates from scanner variations, where a standard deviation of ∼10 gr/l was found, but with similar systemic errors as in experiment C3.

TABLE I. Intrinsic sources of error. Progression estimation errors (gr/l, mean ± std). The experiments marked with an asterisk were manually initialized within the capture range, since the rotation was too large, and the cylinder phantom is rotation invariant.

Experiment	6°	29°	90°
C1. interpolation	0.03 ± 0.2	0.03 ± 0.2	0.08 ± 0.3
C2. registration	0.00 ± 0.0	0.00 ± 0.0*	0.00 ± 0.0*
C3. C1 + C2	0.36 ± 1.0	0.40 ± 1.1	0.25 ± 0.7
C4. C1 + C2 + CT scanner		0.38 ± 9.7	

3.D. Lung model evaluation: Phantom study

Volume correction was validated using the cylinder phantom, with and without progression.

3.D.1. Without progression

The volume was modified by moving the piston. Obviously, this volume change has a marked impact on the density, and models should be able to deal with that. We acquired a scan at $V_b = 1100$ ml and one at $V_f = 1300$ ml. The slope was measured using a third scan of the phantom at a volume $V_b^s = 1000$ ml. Since the mass did not change between scans, the ground truth is that there is no progression anywhere. Nonrigid registration was performed using NC as a similarity measure, in combination with the Euclidean distance metric[38] using three manually annotated corresponding anchor points, in order to deal with the lower contrast in the phantom compared to clinical CT data.

The measures of progression, according to the four models, were computed after registration, and the difference with the known (zero) progression (the error) is given in Table II. When no volume correction was applied (uncorrected method), the systematic progression error deviated substantially from zero (−10.5 gr/l). Both the sponge (0.99 gr/l) and slope models (0.19 and −0.15 gr/l) showed a systematic progression error much smaller than the uncorrected model. The adapted slope models had the smallest systematic error.

3.D.2. With local progression

In order to measure the random and systematic errors of the proposed methods, both the volume was changed physically and the local density of the phantom was adjusted digitally as follows. The middle volume (1100 ml) was chosen as baseline I_b, larger volumes as followup I_f, and smaller volumes were used for slope estimation I_b^s. For a first experiment, we chose I_f at 1180 ml and I_b^s at 1040 ml (∼80 ml difference), and for a second I_f at 1300 ml and I_b^s at 880 ml (∼200 ml difference). A piece of foam was segmented manually, and its density was changed digitally by subtracting the values $v \in \{1, 2, 5, 10, 15, 20, 40\}$ gr/l from the piece of foam. The piece of foam moves and deforms when moving the piston. The ground truth is zero progression in the background and v in the piece of foam. Nonrigid registration was performed using NC as a similarity measure.

TABLE II. Phantom validation, no induced progression. Progression estimation errors (gr/l, mean ± std, absolute mean).

Model	Progression error [gr/l]			
Uncorrected	−10.46	±	17.6	14.4
Sponge	0.99	±	16.1	11.1
Global slope	0.19	±	16.1	11.1
Local slope	−0.15	±	18.2	13.1

The results are shown using bar charts in Fig. 7, where the whiskers denote the standard deviation. The median estimated progression error in the background was zero for all models that apply volume correction, but much larger for the uncorrected model. The median progression in the piece of foam was close to the imposed progression (i.e., a small systematic estimation error). For the 200 ml difference in volume, the accuracy of the sponge model was worse than that of the adapted slope models. Standard errors are ∼10 gr/l for the 80 ml difference and increased to ∼20 gr/l for the 200 ml difference. An example progression image of the phantom is shown in Figs. 10(a) and 10(c).

3.E. Lung model evaluation: Patient data

3.E.1. Linearity

The assumption that the volume-density relation is log-log linear over the breathing cycle, as mentioned in Sec. 2.C, was validated by measuring the volume-density relation in our patient group. CT data at many stages of the breathing cycle are not available due to dose restrictions. Each patient was, however, scanned at two different volumes, enabling measurement of the slopes at different volumes. For each patient, we therefore computed the change in lung volume and the change in median lung density, at baseline between ∼FRC and TLC.

Figure 8 shows the relation between volume and density change for the patient data. It clearly shows a linear relation with a fit value of $R^2 = 0.94$, thereby confirming the log-linearity assumption of the adapted slope models.

3.E.2. Globalized-local vs global

The local methods were globalized as described in Sec. 2.C, which enables comparison with a global progression measure derived from the 15th percentile point.[1] This global progression measure was previously proposed and has been used in several clinical trials.[1,25] It employs a mixed-effects regression model with density as outcome and lung volume as covariate.

The globalized local density change is compared to the global measure. The results are shown in a correlation plot in Fig. 9. The correlation between the globalized local methods and the global result were 0.643, 0.855, 0.865, and 0.882 for the uncorrected, sponge, global, and local slope method, respectively. The highest correlation was obtained by the methods with volume correction. Note that the uncorrected model overestimates progression, because the mean lung volume at

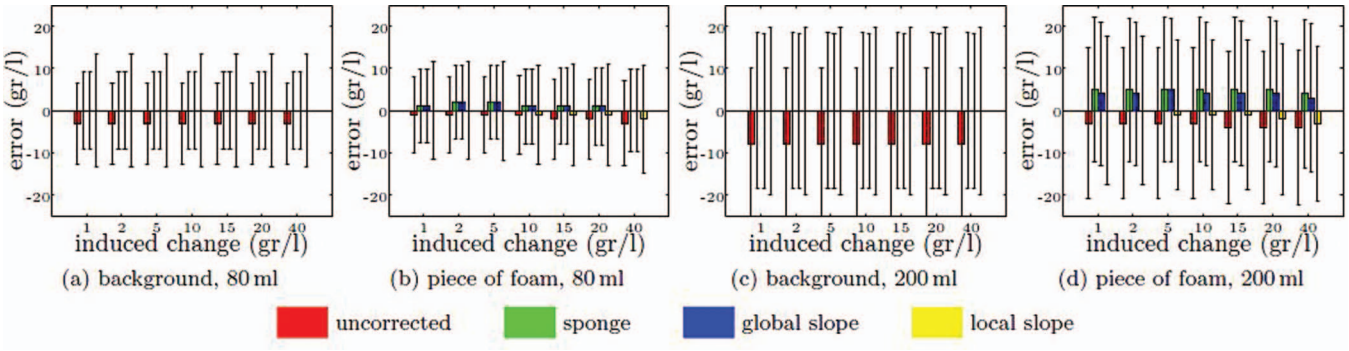

FIG. 7. Phantom study: changing both volume and density. (a) and (b) 80 ml difference, (c) and (d) 200 ml difference, (a) and (c) median background progression, (b) and (d) median progression difference with the ground truth in the piece of foam. (a) Background, 80 ml, (b) piece of foam, 80 ml, (c) background, 200 ml, (d) piece of foam, 200 ml.

baseline was 6.9 l and 7.3 l at follow-up, a difference of 5%. Also note that the globalized local and the global measure are not completely identical, because the global measure was derived from the 15th percentile point of the histograms and the accumulated measure from density differences. Example progression images of patients are shown in Figs. 10(b), 10(d), and 11.

3.E.3. Consistency

Finally, the consistency of the progression estimates was examined by computing the progression between baseline ∼FRC and follow-up TLC and comparing it with the progression estimated using baseline TLC and follow-up TLC (see Fig. 3). Because the baseline images are different between the two trajectories (∼FRC vs TLC), a direct voxel-by-voxel comparison cannot be made. Therefore, we compared the accumulated local progression estimates.

The difference in accumulated local progression between the progression methods as measured by the different estimation trajectories, is -1.03 ± 5.2, -0.90 ± 2.6, -0.85 ± 2.2, and -0.41 ± 2.8 gr/l for the four methods, respectively.

4. DISCUSSION AND CONCLUSIONS

In this study, we present several methods for the estimation of progression of emphysema, based on the registration of follow-up chest CT scans combined with postprocessing.

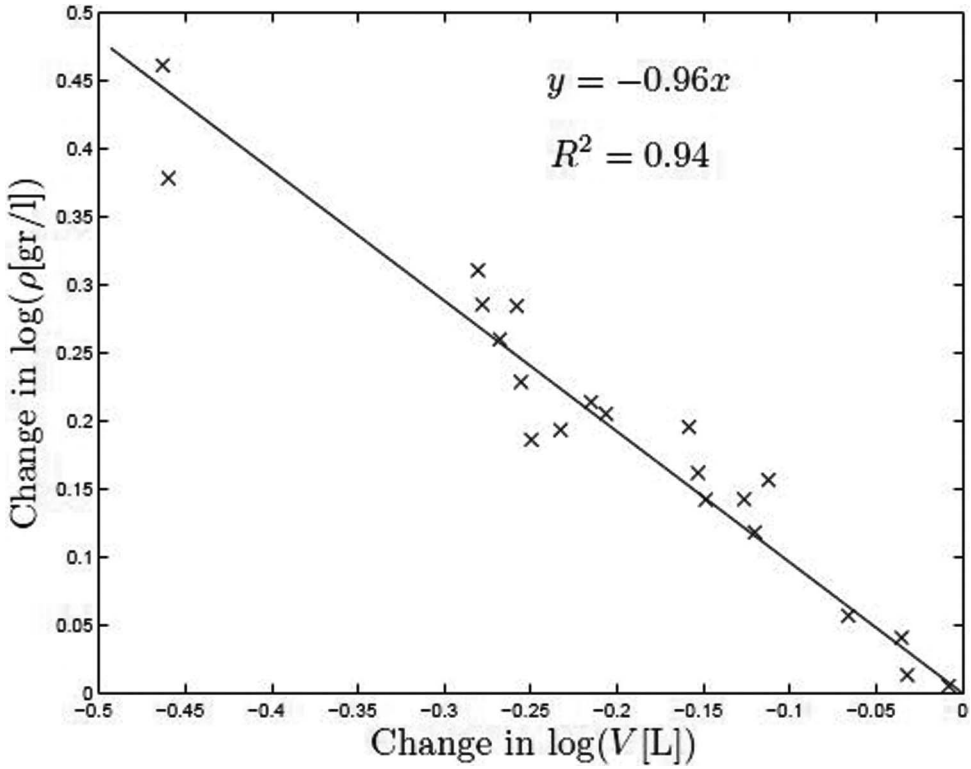

FIG. 8. Linear relationship of the volume-density slope over the breathing cycle. Note that the slope is close to sponge model, and also that it is an average over the patient population opposed to a patient-specific slope.

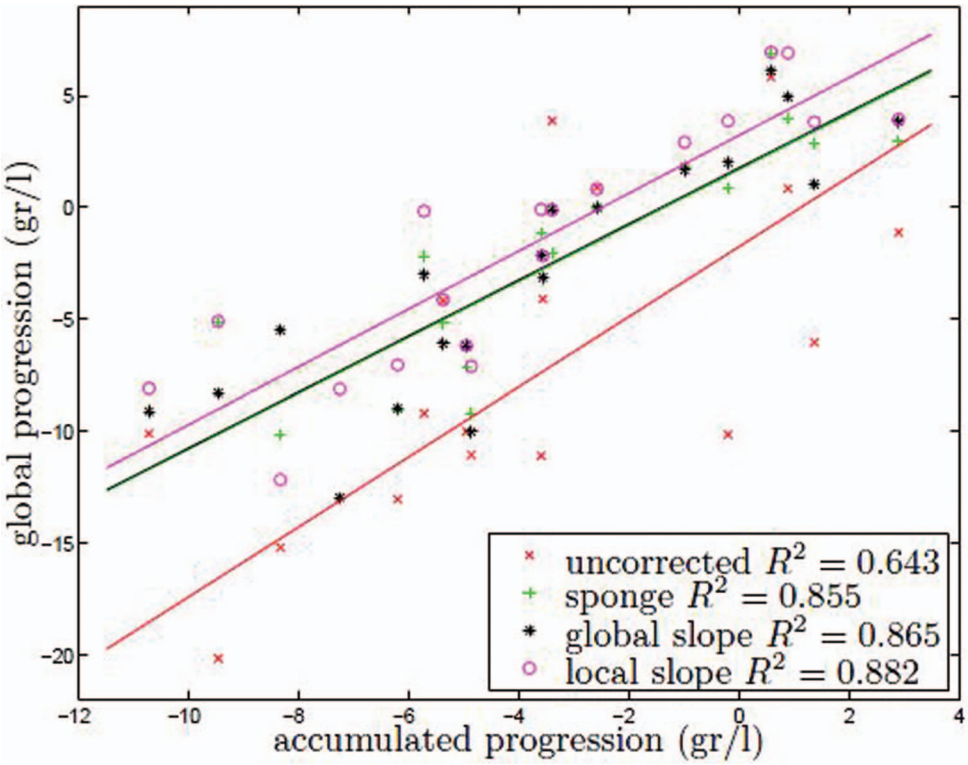

FIG. 9. Correlation plot depicting the accumulated local progression against global progression. There was one outlier of the uncorrected model at (−7.3, −33.0).

Compared to previous methodology,[1] this work additionally offers *local* assessment, and proposes local volume correction with a patient-specific volume-density relation to gain independence of the respiratory state of the lung at image acquisition.

A naïve model that does not perform volume correction was shown to be inaccurate in both phantom and patient data, see Table II and Figs. 7 and 9. Derivation of subsequent methods is therefore based on the assumption that the lung behaves as a dry sponge, i.e., it is mass-preserving. We confirmed that the sponge model does not hold in patient data,[33] see Fig. 2. The deviation between sponge model and observed slopes may be partly explained by changes in lung mass due to changes in blood flow during respiration,[22] and CT scanner related differences between scans.[21] Therefore, a novel adaptation to the sponge model is proposed to accommodate for a global or local patient-specific density-volume slope. A linear relation between the log of the volume and density was assumed, which was confirmed by the experiments. The local volume correction step introduced for the sponge and adapted slope models is based on the spatial Jacobian of the transformation relating baseline with follow-up. Gorbunova *et al.*[15] proposed the use of the Jacobian at the same time,[23] but used it for adapting the registration mechanism in a sponge model setting only, and did not perform local evaluation.

In our study, experiments were performed on phantom and patient data to optimize the registration and verify the lung models. A phantom was therefore constructed that mimics a breathing lung, enabling precise control over volume and density, which is not possible in patients.[21] The constructed phantom additionally enables local validation, also impossible in patients.

The image registration procedure was optimized on CT data of emphysema patients, resulting in a scheme based on normalized correlation, and using lung masks. To our knowledge, this is the first study reporting registration accuracy on thick slice chest CT data of COPD patients, GOLD stage II and III. The registration is accurate with a median registration error of 1.0 mm, which is below half the slice increment.

Two important factors that determine the accuracy of the local lung models are the registration quality and the correction for local lung volume changes (see Table II and Figs. 7 and 9). The accuracy of registration and volume change are related, since from a perfect match a perfect local volume change can be derived. The progression estimation error is therefore largely defined by the residual registration error, and consequently the registration accuracy can act as a partial but important evaluation criterium for the local progression estimation within patients. A full evaluation is, however, not possible, since no local clinical ground truth of emphysema progression exists for patients.

The results of both the phantom and the patient data show that volume correction is needed to evaluate changes in emphysema, and that it is the most important factor. The absence of mass change in the phantom was correctly approximated by the methods that correct for volume. The adapted slope models had the smallest systematic and random error.

When synthetically inducing local density change, the estimation error remains very close to zero in the background for all models that use volume correction. In the modified

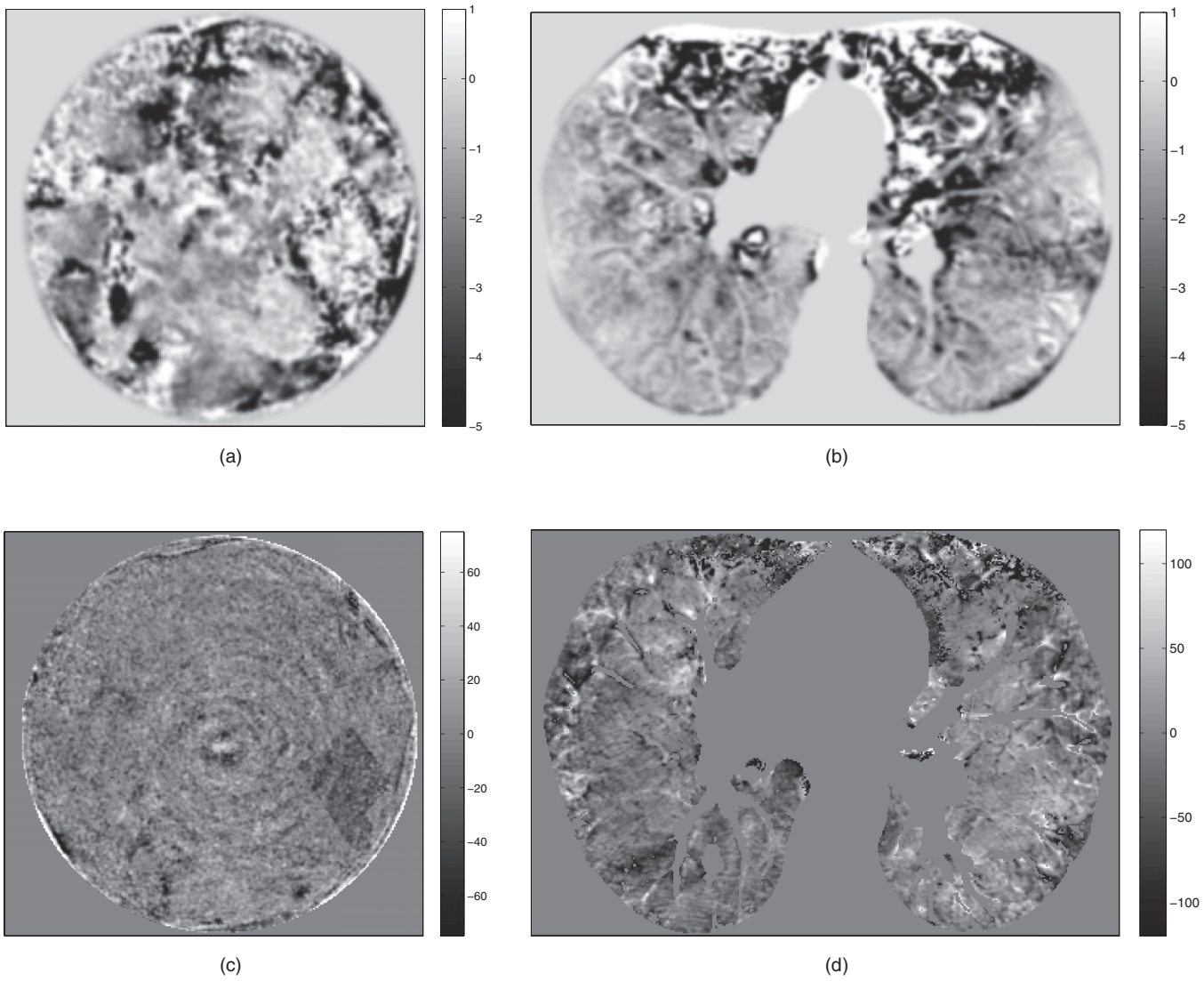

FIG. 10. Slope and progression images. (a) and (c) Phantom, local slope model. The segmented piece of foam was digitally modified by 20 gr/l, and is clearly visible in (c). The circular stripes result from differences in CT image reconstruction. (b) and (d) Patient, local slope model, large vessels are excluded. The ventral part of both the left and right lungs show darkened areas, denoting progression of emphysema. (a) and (b) Local slope $s(\boldsymbol{x})$, (c) and (d) progression, local slope.

area, the systematic error was smaller than 2 gr/l when volume changed by 80 ml (6%), except for the model without volume correction, and increased to 4–5 gr/l for a 200 ml volume difference (19%). In that case, the slope models obtained smaller errors than the sponge model, and therefore more accurately estimated changes in mass. In bullae observed in patient data, the locally adapted slope model is expected to be less accurate, see Fig. 11, second row, third and fourth columns, since the air density in bullae will not change due to respiration. In those areas, a local switch to the sponge or global slope model would possibly improve the accuracy.

The phantom experiments further revealed that the random error for the local models was larger than that of global models. Differences between voxels were noticed up to 100 gr/l, see Fig. 10. Simply rescanning the phantom resulted in standard deviations of 10 gr/l, when summarized over the entire phantom. Evaluation over groups of patients would reduce the standard error to the range of ± 1 gr/l, similar to previous global methods.[39] These experiments suggest the trade-off between locality and standard error, meaning that progression should probably be assessed not on a per-voxel basis, but on somewhat larger areas.

When globalizing the local models, they were consistent with previous global methods. However, the local methods are potentially more sensitive, since small localized changes cannot be found by global tools, and additionally local intensity increases due to, e.g., fibrosis combined with local emphysema progression may go unnoticed using global methods, which may lead to incorrect or incomplete conclusions in drug evaluation trials. We further showed that the local models are consistent in the sense that exchanging the role of the ∼FRC and TLC scan when computing progression does not considerably influence the outcome.

The volume-density slope can also be derived from follow-up scans, possibly resulting in a different estimation. Since this slope is estimated from scans in a different domain, a

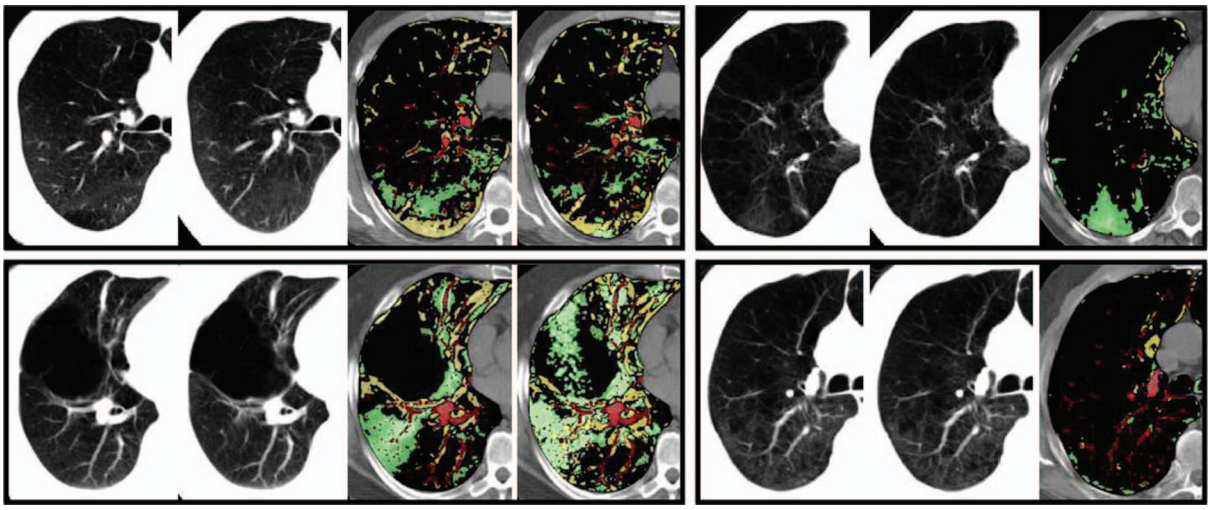

FIG. 11. Example progression maps for four patients. Horizontally shown are the baseline image, registered follow-up (acquired 30 months later) and two (left column) or one (right column) progression maps overlaid on the baseline. For the top left patient, the sponge and local slope model are shown; for the top right patient, the sponge model; for the bottom left patient, sponge and local slope model; and for the bottom right patient, the local slope model. Green and yellow indicate an increase resp. decrease in density, and in red the vessels are indicated, see Sec. 2.D.

resampling step is required using the transformation $\boldsymbol{T}(\boldsymbol{x})$. Alternatively, the slope could be measured at both time instances, after which the mean could be used, or a compensation for the difference could be introduced in the progression model. When comparing two treatment groups, however, it is appropriate to estimate the slope at baseline. During the trial the slope may be affected by the drug or by the disease (by, for example, changes in trapped air), but it is important to measure the overall change between two time points to avoid introducing a bias in the analysis, and to avoid progression being compensated for by adaptation of the slope.

In this study, we used an additional scan to estimate a patient-specific volume-density slope. This comes at the cost of an increase in radiation, which may be a downside in clinical practice. In many drug evaluation trials, however, this is already the standard,[1,25] and a relatively low dose setting is used compared to HRCT. In this study, we used ∼FRC and TLC data, while in some other studies a residual volume scan is taken instead of ∼FRC. Nevertheless, volume correction as derived from the image registration is expected to handle those situations, as long as the registration is successful. The latter is to be expected in the absence of severe pathology, since these methods have been shown to work in a wide variety of cases and lung volume differences, as shown in this paper, but also in, for example, the EMPIRE10 challenge.[13]

This study has some limitations. (1) Patients with severe progression show changes in lung anatomy, which makes it harder to register the CT scans, resulting in larger errors in the progression estimation. (2) A complete clinical validation was not possible due to the lack of a ground truth of local progression. Visual inspection of local changes by an expert is not possible due to the limitations of the human visual system, and computer-based measures (densitometry) should be preferred.[40,41] (3) The phantom CT scan was reconstructed with thin slices, while for the patient data thick slices were used. The goal of the phantom study however was to estimate scanner limitations, while for the patient data we applied settings that are used in practise in clinical trials.[1,25] This subsequently also required a different set of registration parameters. It remains to be investigated if the thick slice protocol which was optimized for global lung densitometry is also optimal for local lung densitometry. (4) The proposed methods are also sensitive to imaging artifacts, such as effects of beam hardening near the ribs, diaphragm, and the shoulders, and streak and motion artifacts around the heart and lung boundary (cardiac, breathing), which show up in the progression images, see Fig. 11. This may give rise to confusion during interpretation. These disturbances however take typical patterns and can be distinguished from real progression. The sensitivity may therefore also be an advantage, since it gives insight in the nature of such artifacts, and these artifacts are also influencing global measurements, which is currently neglected in trials.

In conclusion, the proposed methods that incorporate volume correction to locally estimate differences in lung density (i) effectively eliminate a dependency on lung inspiration level at acquisition time, (ii) accurately predict progression on phantom data, and (iii) are reasonably consistent with global results in patient data. It is therefore a potential tool for assessing local emphysema progression in drug evaluation trials and in clinical practice.

ACKNOWLEDGMENTS

This research was funded by the Dutch Technology Foundation (STW), Grant No. LPG.07998, under the acronym APPEAR (Assessing the Progression of Pulmonary Emphysema by Advanced Registration).

[a]Electronic mail: m.staring@lumc.nl

[1] J. Stolk, H. Putter, M. Bakker, S. Shaker, D. Parr, E. Piitulainen, E. Russi, E. Grebski, A. Dirksen, R. Stockley, J. Reiber, and B. Stoel, "Progression parameters for emphysema: a clinical investigation," Respir. Med. **101**, 1924–1930 (2007).

[2]M. Bakker, H. Putter, J. Stolk, S. Shaker, E. Piitulainen, E. Russi, and B. Stoel, "Assessment of regional progression of pulmonary emphysema with CT densitometry," Chest **134**, 931–937 (2008).
[3]A. Dirksen, J. Dijkman, F. Madsen, B. Stoel, D. Hutchison, C. Ulrik, L. Skovgaard, A. Kok-Jensen, A. Rudolphus, N. Seersholm, H. Vrooman, J. Reiber, N. Hansen, T. Heckscher, K. Viskum, and J. Stolk, "A randomized clinical trial of alpha(1)-antitrypsin augmentation therapy," Am. J. Respir. Crit. Care Med. **160**, 1468–1472 (1999).
[4]S. Hu, E. Hoffman, and J. Reinhardt, "Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images," IEEE Trans. Med. Imaging **20**, 490–498 (2001).
[5]L. Zhang, E. Hoffman, and J. Reinhardt, "Lung lobe segmentation in volumetric X-ray CT images," IEEE Trans. Med. Imaging **25**, 1–16 (2006).
[6]J. Pu, B. Zheng, J. Leader, C. Fuhrman, F. Knollmann, A. Klym, and D. Gur, "Pulmonary lobe segmentation in CT examinations using implicit surface fitting," IEEE Trans. Med. Imaging **28**, 1986–1996 (2009).
[7]E. van Rikxoort, M. Prokop, B. de Hoop, M. Viergever, J. Pluim, and B. van Ginneken, "Automatic segmentation of pulmonary lobes robust against incomplete fissures," IEEE Trans. Med. Imaging **29**, 1286–1296 (2010).
[8]E. van Rikxoort, B. de Hoop, S. van de Vorst, M. Prokop, and B. van Ginneken, "Automatic segmentation of pulmonary segments from volumetric chest CT scans," IEEE Trans. Med. Imaging **28**, 621–630 (2009).
[9]L. Sørensen, S. Shaker, and M. de Bruijne, "Quantitative analysis of pulmonary emphysema using local binary patterns," IEEE Trans. Med. Imaging **29**, 559–569 (2010).
[10]C. Galbán, M. Han, J. Boes, K. Chughtai, C. Meyer, T. Johnson, S. Galbán, A. Rehemtulla, E. Kazerooni, F. Martinez, and B. Ross, "Computed tomography-based biomarker provides unique signature for diagnosis of COPD phenotypes and disease progression," Nat. Med. **18**, 1711–1715 (2012).
[11]B. van Ginneken, K. Murphy, E. van Rikxoort, I. Isgum, B. de Hoop, M. Prokop, P. de Jong, and H. Gietema, in *Annual Meeting of the Radiological Society of North America, Chicago*, 2009.
[12]K. Murphy, B. van Ginneken, E. van Rikxoort, M. de Hoop, J. Pluim, and M. Prokop, in *Proceedings of the Radiological Society of North America, Chicago*, 2008.
[13]K. Murphy, J. Pluim, E. van Rikxoort, P. de Jong, B. de Hoop, H. Gietema, O. Mets, M. de Bruijne, P. Lo, M. Prokop, and B. van Ginneken, "Towards automatic regional analysis of pulmonary function using inspiration and expiration thoracic CT," Med. Phys. **39**, 1650–1662 (2012).
[14]Y. Arzhaeva, M. Prokop, K. Murphy, E. van Rikxoort, P. de Jong, H. Gietema, M. Viergever, and B. van Ginneken, "Automated estimation of progression of interstitial lung disease in CT images," Med. Phys. **37**, 63–73 (2010).
[15]V. Gorbunova, P. Lo, H. Ashraf, A. Dirksen, M. Nielsen, and M. de Bruijne, in *Proceedings of the MICCAI, New York, NY, 2008*, Lecture Notes in Computer Science Vol. 5242 (Springer-Verlag, Berlin Heidelberg, 2008), pp. 863–870.
[16]Y. Yin, E. Hoffman, and C. Lin, "Mass preserving nonrigid registration of CT lung images using cubic B-spline," Med. Phys. **36**, 4213–4222 (2009).
[17]B. Li, G. Christensen, G. McLennan, E. Hoffman, and J. Reinhardt, "Pulmonary CT image registration and warping for tracking tissue deformation during the respiratory cycle through 3-D consistent image registration," Med. Phys. **35**, 5575–5583 (2008).
[18]Y. Yin, E. Hoffman, K. Ding, J. Reinhardt, and C. Lin, "A cubic B-spline-based hybrid registration of lung CT images for a dynamic airway geometric model with large deformation," Phys. Med. Biol. **56**, 203–218 (2011).
[19]V. Gorbunova, S. Durrleman, P. Lo, X. Pennec, and M. de Bruijne, in *Proceedings of the IEEE International Symposium on Biomedical Imaging: Macro to Nano,* Rotterdam, The Netherlands, 2010.
[20]K. Murphy *et al.*, "Evaluation of registration methods on thoracic CT: The EMPIRE10 challenge," IEEE Trans. Med. Imaging **30**, 1901–1920 (2011).
[21]B. Stoel and J. Stolk, "Optimization and standardization of lung densitometry in the assessment of pulmonary emphysema," Invest. Radiol. **39**, 681–688 (2004).
[22]C. Fink, S. Ley, F. Risse, M. Eichinger, J. Zaporozhan, R. Buhmann, M. Puderbach, C. Plathow, and H. Kauczor, "Effect of inspiratory and expiratory breathhold on pulmonary perfusion: Assessment by pulmonary perfusion magnetic resonance imaging," Invest. Radiol. **40**, 72–79 (2005).
[23]M. Staring, M. E. Bakker, D. P. Shamonin, J. Stolk, J. H. C. Reiber, and B. C. Stoel, Towards local estimation of emphysema progression using image registration. In J. P. W. Pluim and B. M. Dawant, editors, SPIE Medical Imaging: Image Processing, volume 7259 of Proceedings of SPIE, page 72590O, Orlando, Florida, 2009.
[24]R. Mull, "Mass estimates by computed tomography: physical density from CT numbers," Am. J. Roentgenol. **143**, 1101–1104 (1984).
[25]J. Stolk, B. Cooper, B. Stoel, A. Rames, O. Rutman, S. Soliman, and R. Stockley, "Retinoid treatment of emphysema in patients on the alpha-1 international registry. the REPAIR study: study design, methodology and quality control of study assessments," Ther. Adv. Respir. Disease **4**, 319–332 (2010).
[26]B. Stoel, H. Vrooman, J. Stolk, and J. Reiber, "Sources of error in lung densitometry with CT," Invest. Radiol. **34**, 303–309 (1999).
[27]S. Klein, M. Staring, K. Murphy, M. Viergever, and J. Pluim, "elastix: a toolbox for intensity-based medical image registration," IEEE Trans. Med. Imaging **29**, 196–205 (2010).
[28]D. Rueckert, L. Sonoda, C. Hayes, D. Hill, M. Leach, and D. Hawkes, "Nonrigid registration using free-form deformations: Application to breast MR images," IEEE Trans. Med. Imaging **18**, 712–721 (1999).
[29]S. Klein, U. van der Heide, I. Lips, M. van Vulpen, M. Staring, and J. Pluim, "Automatic segmentation of the prostate in 3D MR images by atlas matching using localized mutual information," Med. Phys. **35**, 1407–1417 (2008).
[30]S. Klein, J. Pluim, M. Staring, and M. Viergever, "Adaptive stochastic gradient descent optimisation for image registration," Int. J. Comput. Vis. **81**, 227–239 (2009).
[31]J. Reinhardt, K. Ding, K. Cao, G. Christensen, E. Hoffman, and S. Bodas, "Registration-based estimates of local lung tissue expansion compared to xenon-CT measures of specific ventilation," Med. Image Anal. **12**, 752–763 (2008).
[32]K. Ding, J. Bayouth, J. Buatti, G. Christensen, and J. Reinhardt, "4DCT-based measurement of changes in pulmonary function following a course of radiation therapy," Med. Phys. **37**, 1261–1272 (2010).
[33]B. Stoel, H. Putter, M. Bakker, A. Dirksen, R. Stockley, E. Piitulainen, E. Russo, D. Parr, S. Shaker, J. Reiber, and J. Stolk, "Volume correction in computed tomography densitometry for follow-up studies on pulmonary emphysema," Proc. Am. Thoracic Soc. **5**, 919–924 (2008).
[34]B. Stoel, M. Jobse, J. Stolk, and J. Reiber, "Computer aided comparison of CT images in patients with pulmonary emphysema," Eur. Respir. J. **22**, 532s (2003).
[35]C. Xiao, M. Staring, D. Shamonin, J. Reiber, J. Stolk, and B. Stoel, "A strain energy filter for 3D vessel enhancement with application to pulmonary CT images," Med. Image Anal. **15**, 112–124 (2011).
[36]B. Stoel, M. Bakker, J. Stolk, A. Dirksen, R. Stockley, E. Piitulainen, E. Russi, and J. Reiber, "Comparison of the sensitivities of 5 different computed tomography scanners for the assessment of the progression of pulmonary emphysema: a phantom study," Invest. Radiol. **39**, 1–7 (2004).
[37]K. Murphy, B. van Ginneken, S. Klein, M. Staring, B. J. de Hoop, M. Viergever, and J. Pluim, "Semi-automatic construction of reference standards for evaluation of image registration," Med. Image Anal. **15**, 71–84 (2011).
[38]M. Baiker, M. Staring, C. Löwik, J. Reiber, and B. Lelieveldt, in *Proceedings of the MICCAI, Toronto, Canada, 2011*, LNCS Vol. 6892 (Springer Verlag, Berlin Heidelberg, 2011), pp. 516–523.
[39]M. Bakker, J. Stolk, H. Putter, S. Shaker, D. Parr, E. Piitulainen, E. Russi, A. Dirksen, R. Stockley, J. Reiber, and B. Stoel, "Variability in densitometric assessment of pulmonary emphysema with computed tomography," Invest. Radiol. **40**, 777–783 (2005).
[40]E. Cavigli, G. Camiciottoli, S. Diciotti, I. Orlandi, C. Spinelli, E. Meoni, L. Grassi, C. Farfalla, M. Pistolesi, F. Falaschi, and M. Mascalchi, "Whole-lung densitometry versus visual assessment of emphysema," Eur. Radiol. **19**, 1686–1692 (2009).
[41]M. Mascalchi, S. Diciotti, N. Sverzellati, G. Camiciottoli, C. Ciccotosto, F. Falaschi, and M. Zompatori, "Low agreement of visual rating for detailed quantification of pulmonary emphysema in whole-lung CT," Acta Radiol. **53**, 53–60 (2012).