On segmentation of lung parenchyma in quantitative computed tomography of the lung

Gerrit J. Kemerink,[a] Rob J. S. Lamers, Bas J. Pellis, Han H. Kruize, and J. M. A. van Engelshoven
Department of Radiology, University Hospital Maastricht, P. Debijelaan 25, P.O. Box 5800, 6202 AZ Maastricht, The Netherlands

(Received 13 April 1998; accepted for publication 30 September 1998)

Our purpose in this study was to investigate the influence of segmentation threshold and number of erosions on parameters used in quantitative computed tomography (CT) of the lung (erosions are shrink operations on the segmented area). Parameters assessed were mean lung density, area of the segmented lung, two percentiles, and the pixel index, which is the relative area of the histogram below −905 Hounsfield Units (HU). We analyzed images of ten emphysematous and ten nonemphysematous patients, that had been scanned at carina level in inspiration and expiration, using sections of 1, 2, 3, 5, and 10 mm in combination with a standard, a smooth, and an ultrasmooth reconstruction kernel. The lungs were segmented using pixel tracing at thresholds of −200, −400, and −600 HU with 0–4 erosions, followed by histogram analysis. The area of the segmented lungs decreased with 0.9%–3.2% per 100 HU decrease in threshold and with 2.2%–3.1% per erosion, dependent on patient group and respiratory status. Estimated mean lung density changed up to 30% by changing the threshold and the number of erosions. The pixel index and the 10th percentile depended only slightly on threshold and number of erosions, but the 90th percentile showed a strong dependence of up to 40%. It is concluded that the segmentation protocol can have a large impact on densitometric parameters and that standardization is mandatory for obtaining comparable results. Ideally a threshold equal to the average of the densities of lung and soft tissue should be used, but −400 HU will do in a limited but common density range (−910 to −790 HU). For densitometry two erosions are recommended, for volumetry zero erosions should be used. [S0094-2405(98)01912-9]

Key words: densitometry, lung density, CT technology

I. INTRODUCTION

Densitometry of the lung has found an application without sufficient attention to methodological aspects. For instance, CT number calibration was seldomly addressed, while for several scanners considerable errors at low densities can exist.[1] And although the scan protocol may affect densitometric parameters, not all relevant properties of the scan technique were usually specified. Suboptimal protocols have also been applied. It was recently shown that density resolution is an important parameter,[2–4] and that it is quite bad for the so-called thin-section densitometry protocol, which was widely used, also by ourselves.[5] Estimation of parameters that depend on the shape of the lung histogram, like the pixel index and percentiles, should not be done with this protocol.[3,4] Also, different segmentation protocols have been applied, without an assessment of the effect on densitometric results. The consequences of this lack of attention to methodology are suboptimal work and incomparability of most studies.

In an effort to contribute to the required methodological knowledge, in this study we address segmentation of the lung. According to the literature, a few different methods have been applied for the semiautomatic segmentation of the lungs.[6–10] Hedlund *et al.* used gradients in CT numbers together with CT number ranges, both in an edge tracking method and a method that was based on the selection of contiguous pixels within the lung.[6] Kalender *et al.* introduced pixel tracing, whereby contiguous pixels in a certain CT number range were traced along the border of the lung.[7] Other investigators applied methods that binarize the CT image using a threshold that discriminates the lung from soft tissue.[9,10] When segmentation errors occur, they can be corrected manually, or, as shown by Kalender *et al.*[7] and Zagers *et al.*,[8] by software that still may need some manual help.

A serious problem of segmentation is the inadvertent inclusion of pixels partially corresponding to soft tissue surrounding the lung. Such pixels can have a large impact on densitometric parameters of the parenchyma, especially in patients with lungs of low density as in emphysema. Although it is clear that the magnitude of the effect will depend on the segmentation threshold and the number of erosions, this dependence has not yet been studied. According to the literature a few segmentation thresholds have been used: −200 HU,[7] −380 HU,[8] and −400 HU.[10] In several studies the segmentation procedure was poorly described and it was often not reported whether subsequent erosions had been applied. Even the manual of the widely used software package PULMO CT does not specify the threshold and the number of erosions.[7]

Unexamined so far is also the influence of the segmenta-

tion threshold and spatial resolution on the area of the segmented lung regions. The area is relevant because it is the basis for the calculation of the lung volume, which may have important applications in checking and correcting spirometry in longitudinal CT studies, and in quantifying the effect of reduction surgery. Note that spatial resolution can be varied over a wide range with section thickness and reconstruction kernel. Most studies published so far reported section thickness, but none the in-plane spatial resolution, although this latter parameter can have a significant effect on densitometric quantities.[3]

In this study we investigate, for different patients and scan protocols, the effects of segmentation threshold and the number of erosions on several frequently used densitometric parameters, including the area of the segmented lung. We applied a pixel tracing algorithm for automatic segmentation.[11]

II. MATERIAL AND METHODS

The following acquisition parameters were used on our CT scanner (Somatom Plus, Siemens, Erlangen, Germany): a nonspiral scan of 360°, a scan time of 1 s, a high voltage of 137 kV, a tube load of 220 mAs, the large focal spot (1.3×1.2 mm^2), a 512×512 matrix, and a field of view of approximately 310 mm (zoom factor ≈ 1.6) corresponding to a pixel size of about 0.6 mm. Air calibrations were performed daily. The calibration of this scanner at low densities is adequate and stable over time.[1]

Two groups of patients were analyzed: one group consisted of ten persons with emphysema, the second group of ten nonemphysematous persons with various forms of lung disease or normal lungs. The study was approved by the medical ethical committee of our institution and all patients gave informed consent. The two groups were scanned in inspiration and expiration, yielding four categories of data. The scans were made at the level of the carina with the patient in a supine position, using section thicknesses (S) of 1, 2, 3, 5, and 10 mm. The carina level was chosen because segmentation at this position can have some problems. The complete scan protocol causes an effective dose to the patient of approximately 1 mSv. All scan data were reconstructed with three different reconstruction kernels: standard (STD, Siemens code AB7055), soft (SFT, AB7057), and soft detail (SD, AB7059). Consequently, each of the four categories consisted of 150 images. According to the CT manufacturer the full width at half-maximum (FWHM) of the point spread function (PSF) is 1.13 mm for the STD, 1.40 mm for the SFT, and 1.86 mm for the SD kernel. For patient comfort we did not apply spirometry during the ten scans, but we have excluded from analysis those patients whose mean density, averaged over all section thicknesses, had a large standard deviation (see below).

All images were transferred to an image processing station (ICON Power PC, Siemens Gammasonics, Hoffman Estate, IL). Using pixel tracing,[7,11] the lungs were semiautomatically segmented using three different thresholds (T): -200 HU, -400 HU, and -600 HU. A starting point for segmentation was automatically determined by software. Values of -200 and -400 HU have been used before according to the literature; we added -600 HU to see the effect of a low threshold. The tracing algorithm yields a lung contour consisting of contiguous pixels with CT numbers as close as possible to the threshold, but not exceeding this. When user interaction was required, the action was logged to get an impression of the frequency of the problem in relation to the segmentation threshold and scan parameters. Subsequently, both lung regions were subjected to 0, 1, 2, 3, or 4 erosions (E). One erosion corresponds to the removal of all pixels that lay along the circumference of the segmented area, and that have at least one of their eight neighbors outside the area. Histograms were created from the left lung, the right lung, and the total lung, and the following histogram parameters were calculated: the area of the lung region, the mean CT number, the pixel index PI(-905), the 10th percentile $P(10)$, and the 90th percentile $P(90)$. The mean CT number is for the (water-equivalent) lung directly related to the mean density (ρ): $\rho = (1000 + \text{CT number})/1000$ g/cm^3. Because of this simple relation we will also present density sometimes in terms of the CT number. The pixel index PI(-905) defines the relative area of the histogram below -905 HU, and $P(10)$ and $P(90)$ define the CT number below which 10%, respectively 90%, of the histogram area extends. The pixel index (or density mask) has been used as a quantitative measure for emphysema.[9] We chose PI(-905) as the average of -900 and -910 HU, which both have been used (Refs. 9 and 12 and many subsequent studies).

A problem when trying to illustrate the systematic effects of segmentation threshold, number of erosions, section thickness, and reconstruction kernel, is the variation in histogram parameters introduced by two other mechanisms. First, scans with different section thickness necessarily differ in the (heterogeneous) lung tissue that is sampled. Second, small differences in the degree of inspiration or expiration will exist. To reduce the impact of these two effects we included in our analysis only patients whose standard deviation of mean density, calculated from all sections in each single respiratory status, was less than 6 HU for emphysematous patients, and less than 12 HU for nonemphysematous patients ($T = -400$ HU, $E = 2$). The density is thus taken as a criterion to judge the constancy of inspiration and the identicalness of the tissue studied.

A further reduction was obtained by averaging values from all patients within each of the four categories. This was done after largely eliminating interpatient differences in lung properties by subtracting from each histogram parameter a patient specific reference value. This reference value was the average over all 15 combinations of section thickness and reconstruction kernel of that same parameter for $T = -400$ HU and two erosions of that patient. Thus, each patient has for each respiratory status and each histogram parameter a reference value. Averaging over 15 combinations was performed to increase the robustness of the reference value. Note that the present approach assumes that a densitometric parameter changes for all patients in a similar way around his reference value when segmentation threshold, number of erosions, or spatial resolution are changed.

TABLE I. Frequency of manual operations required to correct automatic segmentation.[a]

Threshold (HU)	Inspiration			Expiration		
	−200 (%)	−400 (%)	−600 (%)	−200 (%)	−400 (%)	−600 (%)
Emphysema						
Separation R&L lungs	61	57	30	68	57	24
Exclusion R bronchus	85	53	15	97	69	14
Exclusion L bronchus	17	7	2	6	2	2
Nonemphysema						
Separation R&L lungs	37	27	13	28	18	10
Exclusion R bronchus	78	62	17	73	29	2
Exclusion L bronchus	2	1	0	29	20	0

[a]Each percentage pertains to 150 images: 10 patients, 5 section thicknesses, and 3 reconstruction kernels. CT images were taken at the carina level.

Lung areas of each patient were expressed as a percentage of his reference area, and changes in area are given in percent points (i.e., as simple differences between these percentages).

The advantage of averaging a histogram parameter over all patients in a given category is that trends are much more clearly reflected than they are in the data from a single patient. Therefore, this average was further used in regression analysis. Another approach to examining the large set of data would have been to perform regression analysis on the scatter plots of individual patients, and to average the slopes of all patients belonging to the same category. With this latter method essentially identical results were obtained, but less suitable graphs for visual presentation.

According to simulations to be discussed, it is expected that for a given segmentation threshold the area is approximately linearly dependent on spatial resolution. As an effective measure for spatial resolution we used the geometric mean $R=(\mathrm{FWHM}^2*S)^{1/3}$, with FWHM the width of the PSF and S the section thickness. This particular choice for R accounts for the fact that partial volume averaging occurs in all three orthogonal directions. This R resulted in an approximately linear relation with lung area, and thus in a meaningful presentation of the data of all 15 combinations of kernel and section thickness in a single graph. The same manner of presentation was chosen for ρ, PI(-905), $P(10)$, and $P(90)$, because also rather good linearity was obtained for these parameters. For PI(-905) the quality of a linear fit, as judged from χ^2 and visual inspection, was considerably better when R was used instead of the previously used inverse square root of the sample volume,[3] while for $P(10)$ and $P(90)$ the quality of the fit was very similar. For consistency, R was used here in all cases.

The dependence of the size of the segmented area on different reconstruction kernels and segmentation thresholds was also studied theoretically. We assumed a sharp density step between the lung and the surrounding soft tissue, and convolved this step with the point spread function of the reconstruction kernels. For simplicity we assumed in most simulations that the thorax cavity locally has flat walls and that only the in-plane spatial resolution mattered. The results show the CT number-dependent shift in the apparent lung border with respect to the original (step) position.

We also measured the volume of a phantom that consisted of two polyethene bottles that had been cast into polymethylmethacrylate (PMMA or perspex). The volume of each compartment was determined by filling with water and weighing. The phantom was scanned with its long axis making angles of about 10° with the axis of the scanner, both in the horizontal and in the vertical plane. The reason for this way of positioning was to get appropriate delineation of the flat bottom and top of the bottles in transversal images. Different section thicknesses, reconstruction kernels, and interslice distances were applied. The zoom factor was 1.7. The volume was calculated by adding the areas of all scans and multiplying this sum with the interslice distance.

Due to the very large amount of lung histogram related data (we analyzed the right, the left and the total lung in 600 images) only a limited, more or less representative part will be presented. For instance, we only give results for the total lung and not for the two separate lungs. Since densitometry has been applied, in particular, for the quantification of emphysema during inspiration, we will generally present data for this category, but when relevant also data for the other categories will be given.

III. RESULTS

A. Manual correction of automatically generated contour

Table I shows the frequency of manual operations required to correct the automatic segmentation. These numbers are averages over 10 patients and 15 combinations of section thickness and reconstruction kernel (thus 150 images). Figure 1 illustrates the dependence on effective spatial resolution (R): to improve statistics we added all data of the emphysematous and nonemphysematous groups, both in inspiration and expiration, resulting in 40 cases per point. At a threshold of -600 HU the automatic segmentation was in several cases poor for nonemphysematous patients in expiration. This was caused by the relatively high density of the

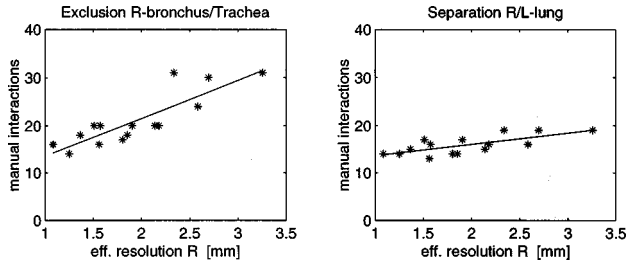

FIG. 1. Manual corrections per 40 cases to improve automatic segmentation of the lungs, as a function of effective spatial resolution $R(=[\mathrm{FWHM}^2_{\mathrm{PSF}}*S]^{1/3})$. Data of emphysematous and nonemphysematous patients in inspiration and expiration were added. The segmentation threshold was -400 HU. Left: the manual exclusion of the trachea or right bronchus. Right: manual separation of the left and right lung.

lung in this category (mean density -795 ± 73 HU, maximum -652 HU), which made that parts of the lung were excluded.

B. Area of segmented lung regions

Figure 2 shows, for the group of emphysematous patients in inspiration, the dependence of the area of the segmented lung on the segmentation threshold, effective spatial resolution, and number of erosions. Note that each point is the average of ten patients, and that each patient's value was expressed as a percentage of his reference area minus 100%. For the other three data categories (emphysema expiration, nonemphysema inspiration, and expiration) the results were qualitatively similar, except for some differences in slopes and intercepts. Table II gives these slopes of the scatter plots

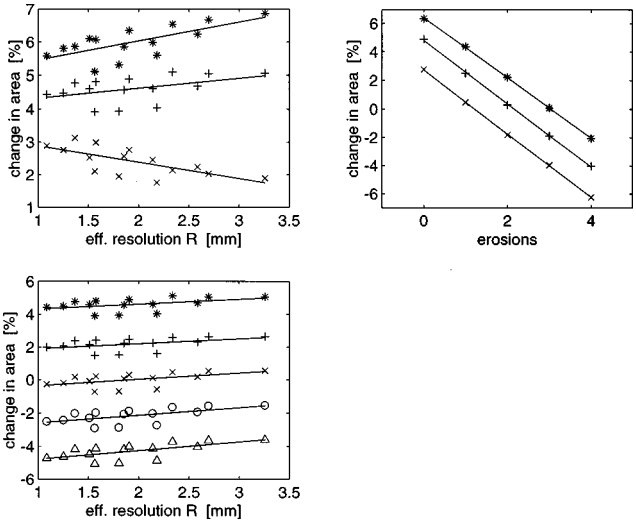

FIG. 2. The percentage change in the area of the total segmented lung for emphysematous patients in inspiration. Data are averages over ten patients. Left top: the change in area as a function of the effective spatial resolution R for zero erosions. (*: $T=-200$ HU, +: $T=-400$ HU, ×: $T=-600$ HU). Right top: the change as a function of the number of erosions for 2 mm sections and the SD kernel. (*: $T=-200$ HU, +: $T=-400$ HU, ×: $T=-600$ HU). Bottom: the change as a function of R and different numbers of erosions for $T=-400$ HU. From top to bottom: from zero to four erosions. The average of the reference areas was 325 ± 54 cm^2.

TABLE II. Slope[a] of the scatter plot of the segmented lung area versus effective spatial resolution.

Threshold (HU)	−200	−400	−600
Emphysema, inspiration	0.57±0.15	0.29±0.15	−0.50±0.13
Emphysema, expiration	1.82±0.32	1.27±0.30	0.47±0.27
Nonemphysema, inspiration	0.36±0.37	−0.38±0.34	−1.69±0.37
Nonemphysema, expiration	2.80±0.54	1.78±0.53	−0.39±0.57

[a]The unit of the slope is a percent of reference area per unit change of effective spatial resolution (%/mm). Note: no erosions have been applied.

of segmented area versus effective resolution R for each of the four data categories, all three thresholds, and zero erosions.

The dependence of the area on the segmentation threshold was calculated from the difference in areas obtained for the thresholds of -200 HU and -600 HU. For zero erosions one finds: $-0.9\pm0.2\%$ of the reference area per 100 HU decrease in threshold for the emphysematous patients in inspiration, $-1.2\pm0.2\%/100$ HU for this group in expiration, $-1.6\pm0.3\%/100$ HU for the nonemphysematous patients in inspiration, and $-3.2\pm0.6\%/100$ HU in expiration. The change in segmented area per erosion, for 2 mm thick sections reconstructed with the SD kernel (an average R) and averaged over the three different thresholds, is for the emphysematous patients in inspiration $-2.2\pm0.1\%$ and in expiration $-2.5\pm0.1\%$, and for nonemphysematous patients, respectively, $-2.4\pm0.1\%$ and $-3.1\pm0.5\%$. These results show that one erosion already has a large impact on the estimated lung area, and consequently on the lung volume, if that would be derived.

C. Mean lung density

Figure 3 shows, for the emphysematous patients in inspiration, the dependence of the mean CT number on the segmentation threshold, effective spatial resolution, and number of erosions. Large differences in mean density up to 30% can be obtained (the mean CT number for this category of data was -901 HU, corresponding to a density of 0.099 g/cm^3). Results for the other three data categories were similar. The slopes of the scatter plots of the mean CT number versus effective spatial resolution, for $T=-400$ HU and zero, two, and four erosions, are given in Table III.

The decrease of the mean density with decreasing segmentation threshold was calculated from the differences between the mean densities obtained for the thresholds of -200 HU and -600 HU. For two erosions one obtains: -1.8 ± 0.7 HU per 100 HU threshold decrease for the emphysematous patients in inspiration, -2.6 ± 0.9 HU/100 HU for this group in expiration, -3.2 ± 1.2 HU/100 HU for the nonemphysematous patients in inspiration, and -7.0 ± 1.2 HU/100 HU in expiration.

D. PI(−905), *P*(10), and *P*(90)

Figures 4 and 5 illustrate the dependence of the pixel index PI(-905) and the percentile $P(10)$ on the segmentation threshold, spatial resolution, and number of erosions,

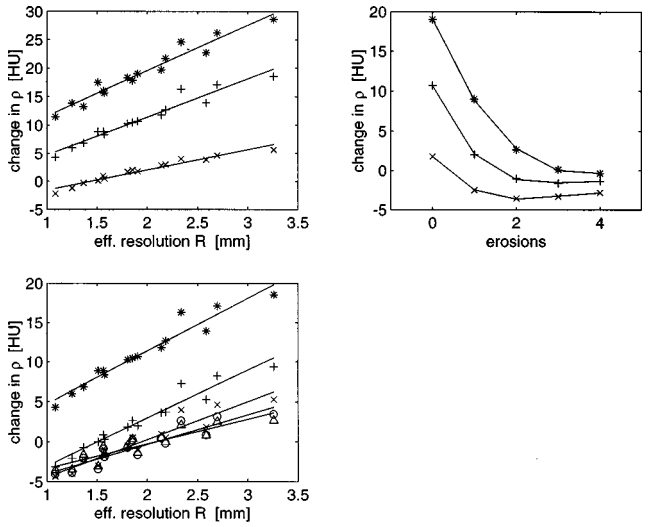

FIG. 3. The change in the mean CT number of the total segmented lung for emphysematous patients in inspiration. Data are averages over ten patients. Left top: the change in the mean CT number as a function of the effective spatial resolution R for zero erosions. ($*$: $T=-200$ HU, $+$: $T=-400$ HU, $\times$: $T=-600$ HU). Right top: the change as a function of the number of erosions for 2 mm sections and the SD kernel. ($*$: $T=-200$ HU, $+$: $T=-400$ HU, $\times$: $T=-600$ HU). Bottom: the change as a function of R and different numbers of erosions for $T=-400$ HU. From top to bottom: from zero to four erosions. The average of the reference values of the mean CT number was -901 ± 14 HU.

again for the emphysema group in inspiration. Both parameters are only slightly dependent on the threshold and number of erosions. The dependence on the spatial resolution is much more pronounced. The same observations hold for the other three data categories.

Figure 6 presents the results for the percentile $P(90)$, also for the emphysema group in inspiration. Changes of more than 40% of the density corresponding to $P(90)$ are possible due to the changing segmentation threshold and number of erosions [the average $P(90)$ is -821 HU for this group, corresponding to a density of 0.179 g/cm^3]. The slope of the scatter plot of $P(90)$ versus R decreases from (on average) positive to negative when the number of erosions is sufficiently increased: for zero erosions and a threshold of -400 HU, the slopes for each of the four categories are 11 HU/mm (emph., insp.), 0 HU/mm (emph., exp.), 8 HU/mm (nonemph., insp.), and -4 HU/mm (nonemph., exp.), while for four erosions the corresponding values are -1, -9, -2, and -12 HU/mm.

TABLE III. Slope of the scatter plot of the mean CT number of the lung versus effective spatial resolution. Note: Segmentation threshold -400 HU.

Erosions	0 (HU/mm)	2 (HU/mm)	4 (HU/mm)
Emphysema, inspiration	6.7 ± 0.4	4.7 ± 0.4	3.1 ± 0.4
Emphysema, expiration	4.0 ± 0.6	2.4 ± 0.5	0.6 ± 0.4
Nonemphysema, inspiration	6.9 ± 0.8	4.9 ± 0.8	2.7 ± 0.8
Nonemphysema, expiration	4.5 ± 0.9	2.9 ± 0.8	0.8 ± 0.7

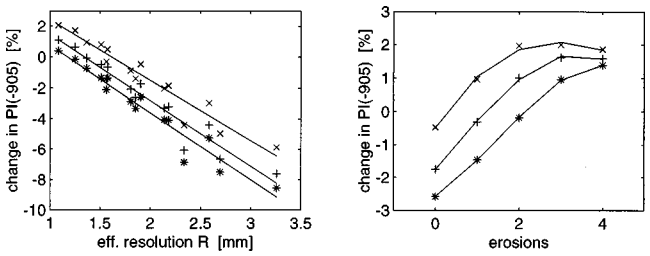

FIG. 4. The change in the pixel index PI(-905) of the total segmented lung for emphysematous patients in inspiration. Data are averages over ten patients. Left: the change in PI(-905) as a function of the spatial resolution for zero erosions. Right: the change in PI(-905) as a function of the number of erosions for 2 mm sections and the SD kernel. ($*$: $T=-200$ HU, $+$: $T=-400$HU, $\times$: $T=-600$ HU). Average of the reference PI(-905) values was $64\pm8\%$.

E. Simulations and phantom study

Figure 7 shows the results of the convolution of the point spread function of the STD and the SD reconstruction kernels with a step function representing the border between the lung and soft tissue. In these simulations we assumed a CT number of lung of -900 HU, the average for our group of emphysematous patients in inspiration, and of 50 HU for soft tissue. The shift for a given threshold is in a first approximation linearly dependent on the spatial resolution, and so is the area for shifts small compared to the linear size of the region, as is the case here. Simulations with cylindrically and spherically shaped borders gave nearly identical results as long as the radii were larger than about one centimeter.

The results of the measurements on the two-bottle phantom are shown in Table IV. We used a segmentation threshold of -430 HU, a value that was the average CT number of air in the bottles and the surrounding PMMA. The error in the estimated volume is small for this theoretically optimal threshold, and independent of spatial resolution.

IV. DISCUSSION

A. Manual correction of automatically generated contours

At the carina level the number of manual corrections that is required is high, and this number increases with worsening

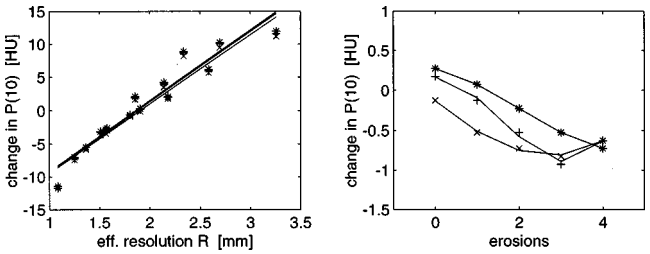

FIG. 5. The change in percentile $P(10)$ of the total segmented lung for emphysematous patients in inspiration. Data are averages over ten patients. Left: the change in $P(10)$ as a function of the spatial resolution for zero erosions. Right: the change in $P(10)$ as a function of the number of erosions for 2 mm sections and the SD kernel. ($*$: $T=-200$ HU, $+$: $T=-400$ HU, $\times$: $T=-600$ HU). The average of reference $P(10)$ values was -974 ± 17 HU.

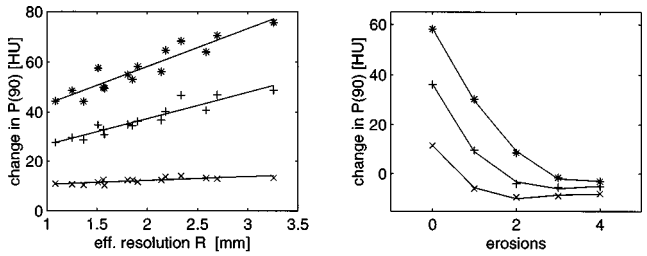

FIG. 6. The change in percentile $P(90)$ of the total segmented lung for emphysematous patients in inspiration. Data are averages over ten patients. Left: the change in $P(90)$ as a function of the spatial resolution for zero erosions. Right: the change in $P(90)$ as a function of the number of erosions for 2 mm sections and the SD kernel. ($*$: $T=-200$ HU, $+$: $T=-400$ HU, $\times$: $T=-600$ HU). The average of the reference $P(90)$ values was -821 ± 15 HU.

of the spatial resolution, but decreases with a decreasing threshold. These findings are in agreement with the assumption that thin septa and the finite spatial resolution, possibly in combination with some motion, are causing the failure during automatic segmentation. Note that poorer spatial resolution enhances the smearing out of small structures, thus lowering the CT values of thin septa. A low threshold is thus better suited for segmentation along thin septa, but a too low value, as our threshold of -600 HU for nonemphysematous patients in expiration, can lead to the exclusion of parts of the lung. At a threshold of -400 HU, manual corrections were, on average, 30% less frequently required than at the more commonly used threshold of -200 HU.[7] We did not yet implement automatic correction algorithms, as, for instance, was done by Kalender *et al.*[7] and Zagers *et al.*[8] Note

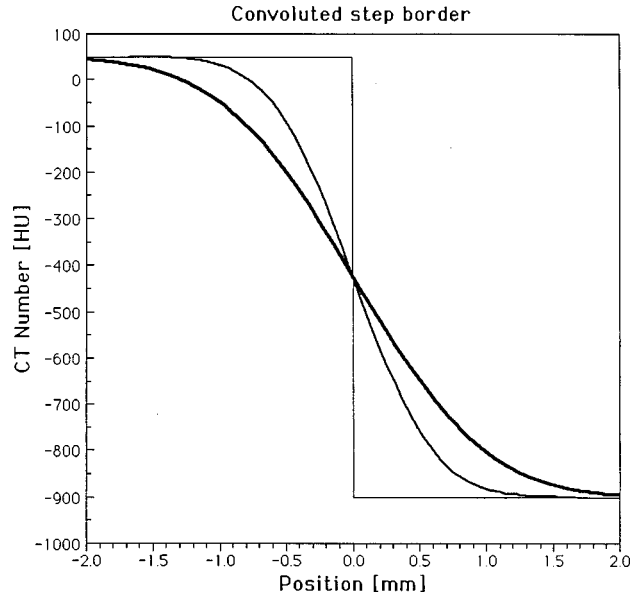

FIG. 7. Simulation of the lung border in a CT image by the convolution of the point spread function of the reconstruction kernel with a step border between the lung (mean CT number -900 HU) and soft tissue (mean CT number 50 HU). Thin line: the border assumed in case of infinitely good spatial resolution. Medium: the border obtained when using the STD kernel ($FWHM_{PSF}=1.13$ mm). Fat: the border obtained when using the SD kernel ($FWHM_{PSF}=1.86$ mm).

TABLE IV. Error in the volume estimate of a two-bottle phantom.[a]

Section thickness (mm)	Slice distance (mm)	Recon. kernel	Error in CT estimate Left (%)	Right (%)	Total (%)
1	10	STD	−0.72	−0.32	−0.52
1	10	SD	−0.69	−0.26	−0.47
5	5	STD	−0.33	−0.35	−0.34
5	5	SD	−0.33	−0.40	−0.37
10	10	STD	−0.01	−0.14	−0.06
10	10	SD	−0.11	−0.05	−0.03

[a]Volumes: left bottle 2.670 *l*, right 2.686 *l*, and total 5.355 ± 0.005 *l*. The segmentation threshold was -430 HU.

that Zagers *et al.* reported that even with an automatic correction between 29% and 48% of the images required user interaction.

B. Area of segmented lung regions

A good starting point for the discussion of the segmented lung area is Fig. 7. It shows that for a threshold that is equal to the average CT number of lung and soft tissue, the position of the border of the lung is independent of the spatial resolution. When a threshold is higher than this average CT number, the border will be found too far outward and the area will be overestimated. If spatial resolution is made worse in this situation, the overestimation will increase. For low thresholds the opposite occurs. Our experimental findings are largely, but not fully, in agreement with the simulations: (1) the segmented area decreases with decreasing threshold, as expected. (2) For all four categories the slope of the area versus the spatial resolution curve decreases with a decreasing segmentation threshold, and the slopes are positive for a threshold of -200 HU, as one would expect. However, for -600 HU the slope is negative only for three out of the four categories, and at -400 HU the difference in slopes between inspiration and expiration is opposite to what one would expect. One obvious flaw of our model is that it neglects medium-sized vasculature that is depicted in the CT images with a broad range of CT numbers, as a consequence of the finite resolution. These vessels can drop below the segmentation threshold when section thickness is increased or in-plane resolution is lowered. This phenomenon leads to an increase in area with poorer spatial resolution that counteracts the decrease at low thresholds that is associated with a (full) step border in other places.

From a theoretical point of view, and neglecting the complications just discussed, it is best to use a threshold that is close to the average CT number of the lung and soft tissue. Our phantom study corroborates this: the volumes derived for the two bottles were nearly identical for the STD and SD kernel, and they matched the true values quite well. At least for this phantom section thickness and reconstruction kernel are not critical for volume estimation. When the threshold deviates from the optimal value, an estimate of the systematic error can be made: a difference of 100 HU between the threshold and mean CT number of the lung and soft tissue will result in an error between 0.9% and 3.2% in the area

estimation, depending on patient and respiratory status. Therefore, if one would decide to use in general practice a threshold of −400 HU, this threshold would, for instance, for lungs with a mean CT number between −910 HU and −790 HU never be off by more than 30 HU from the optimal value. The systematic error in the area would then generally be less than 1%. In this study no mean density below −910 HU was observed, and only in the group of nonemphysematous patients four persons exceeded −790 HU in expiration (the highest was −652 HU). However, for volumetry it seems also possible to make a quick estimate of the mean density and to choose an optimal threshold accordingly. Note that in longitudinal studies random errors are probably more important than threshold related effects, at least as long as the threshold is kept fixed and the same scan protocol is used. Random errors depend among others on the degree of control of inspiration and on patient positioning.

For volumetry no erosions should be applied, because one erosion already reduced the area with about 2%–3%. Note that these percentages are approximately proportional to the linear dimension of a pixel.

C. Mean lung density

Large differences of up to 30% in estimated density could be introduced by changing the segmentation threshold and number of erosions. Note that this number is for the average of ten patients; in individual cases the effect can still be larger. These variations are caused by the inclusion or exclusion of relatively high-density pixels at the circumference of the lung region. Such pixels are the result of spillover of soft tissue into the lung region by the finite resolution. They can be removed by lowering the segmentation threshold and/or applying erosions. For a given threshold, the density decreased only very slowly when more than two erosions were applied. Small differences for different thresholds remained, however, probably due to structures that are above the lower threshold, but below the higher. In case of the higher threshold many erosions may be required in some locations to remove these structures to the degree that they were excluded at the lower threshold. As discussed in the section on area estimation, vasculature may be involved, but also other small structures of high density (e.g., septa), as well as transitions between the lung and soft tissue that appear smooth in the CT image.[6]

In the present results it is shown that the estimated mean density in patient studies is dependent on spatial resolution, i.e., on section thickness and reconstruction kernel, through its effect on segmentation. The dependence is not very strong, but in previous publications we suggested (incorrectly) that there was no relation, which was based on the consideration of bulk materials.[2,3] This dependence is strongest when no erosions are applied, and it is reduced by eroding. For instance, for a threshold of −400 HU and averaging over all four categories of patients, the slope reduced from 5.5 HU/mm to 1.8 HU/mm after four erosions. The slopes are positive because with increasing R, i.e., poorer resolution, the border profile becomes flatter so that for a given threshold more high-density pixels will be included in the segmented region.

For many purposes a threshold of −400 HU and two erosions may be adequate. For lungs with very low or high density, the threshold should be adapted, e.g., to the average CT number of the lung and soft tissue, as discussed above. An advantage of using −400 HU as a threshold, as compared to the more frequently used −200 HU, is that the chance for successful automatic segmentation is somewhat higher. The rather widely used program PULMO CT applies values that generally will be nonoptimal: a threshold of −200 HU and one erosion.[7]

When comparability between different investigations is mandatory, one should use the same segmentation parameters, and preferably also the same scan protocol.

D. PI(−905), $P(10)$, and $P(90)$

The pixel index PI(−905) and the percentile $P(10)$ are only slightly dependent on the segmentation threshold and number of erosions. Changing the segmentation threshold and number of erosions affects initially only a limited number of pixels that belong to the high-density part of the CT number histogram of the lung (well above −905 HU). Consequently, the relative change in PI(−905) is roughly equal to the relative change in the histogram area. For $P(10)$, when properly expressed in density, the relative change is still smaller. However, both parameters are rather strongly dependent on the section thickness and reconstruction kernel (i.e., on R or the sample volume).[3]

$P(90)$ depends heavily on the segmentation threshold and number of erosions, and to a lesser degree on spatial resolution. Changes of up to 40% in the density corresponding to $P(90)$ are possible because high-density pixels are involved.

The effect of decreasing resolution for uniform cellular materials is a smoother image, a narrower histogram, and thus a lower $P(90)$.[3] For a segmented lung this effect [a negative slope of $P(90)$ versus R] becomes only manifest after the removal of high-density pixels at the circumference of the lung region by a sufficiently low threshold and enough erosions.

The dependence of PI(−905), $P(10)$, $P(90)$ on resolution was studied before. For the quantification of these parameters it was recommended to use a section thickness and reconstruction kernel with $S*(1.064*\mathrm{FWHM})^2 \geqslant 8\ \mathrm{mm}^3$, and for comparability, to adhere to a standard scan protocol.[2,3]

V. CONCLUSIONS

Most clinical densitometric studies published so far paid little attention to methodological aspects. This has led to work that was not always optimal, and to incomparability of different studies. Only with due attention to scan and analysis protocols quantitative CT of the lung has a chance of a more general and useful application.

In this study we investigated, for emphysematous and nonemphysematous patients and for different scan protocols, the influence of the segmentation threshold and number of erosions on the following densitometric parameters: area of

the segmented lung, mean lung density, the pixel index PI(-905), and the percentiles $P(10)$ and $P(90)$. The results indicate that different scan and segmentation protocols, in principle, can lead to unacceptably large differences in densitometric parameters. Ideally a segmentation threshold equal to the average of the densities of the lung and soft tissue should be used. However, for much of the work a threshold of -400 HU will be adequate. For densitometry two erosions can be recommended, while for volumetry zero erosions should be applied. For comparability of data the same scan and analysis protocols should be applied.

ACKNOWLEDGMENTS

The authors thank Marthijs Buursink for assistance with preliminary experiments, Geert Wijnhoven for help with the acquisition of patient data, and A. J. T. Zwartkruis (Philips Medical Systems, Best, Netherlands) for his comment on the manuscript.

[a]Correspondence: Gerrit J. Kemerink, Ph.D., Department of Radiology, University Hospital Maastricht, P.O. Box 5800, 6202 AZ Maastricht, The Netherlands. Tel: 31/43/3876910; Fax: 31/43/3876909; Electronic mail: gkk@ms-azm-1.azm.nl

[1]G. J. Kemerink, R. J. S. Lamers, G. R. P. Thelissen, and J. M. A. van Engelshoven, ''Scanner conformity in CT densitometry of the lungs,'' Radiology **197**, 749–752 (1995).

[2]G. J. Kemerink, H. H. Kruize, R. J. S. Lamers, and J. M. A. van Engelshoven, ''Density resolution in quantitative computed tomography of foam and lung,'' Med. Phys. **23**, 1697–1708 (1996).

[3]G. J. Kemerink, H. H. Kruize, R. J. S. Lamers, and J. M. A. van Engelshoven, ''CT lung densitometry: dependence of CT numbers histograms on sample volume and consequences for scan protocol comparability,'' J. Comput. Assist. Tomogr. **21**, 948–954 (1997).

[4]G. J. Kemerink, H. H. Kruize, and R. J. S. Lamers, ''The CT's sample volume as an approximate, instrumental measure for density resolution in densitometry of the lung,'' Med. Phys. **24**, 1615–1620 (1997).

[5]R. J. Lamers, G. R. Thelissen, A. G. Kessels, E. F. Wouters, and J. M. van Engelshoven, ''Chronic obstructive pulmonary disease: evaluation with spirometrically controlled CT lung densitometry,'' Radiology **193**, 109–113 (1994).

[6]L. W. Hedlund, R. F. Anderson, P. L. Goulding, J. W. Beck, E. L. Effmann, and C. E. Putman, ''Two methods for isolating the lung area of a CT scan for density information,'' Radiology **144**, 353–357 (1982).

[7]W. A. Kalender, H. Fichte, W. Bautz, and M. Skalej, ''Semiautomatic evaluation procedures for quantitative CT of the lung,'' J. Comput. Assist. Tomogr. **15**, 248–255 (1991). This work is the basis for the software package PULMO CT (Siemens, Erlangen, Germany). Values for threshold and erosions were received upon request: $T=-200$ HU, 1 erosion.

[8]R. Zagers, H. A. Vrooman, N. J. M. Aarts, J. Stolk, L. J. Schultze Kool, A. E. van Voorthuisen, and J. H. C. Reiber, ''Quantitative analysis of computed tomography scans of the lungs for the diagnosis of pulmonary emphysema,'' Invest. Radiol. **30**, 552–562 (1995).

[9]N. L. Müller, C. A. Staples, R. R. Miller, and R. T. Abboud, '' 'Density Mask': an objective method to quantitate emphysema using computed tomography,'' Chest **94**, 782–787 (1988).

[10]K. T. Bae, R. M. Slone, D. S. Gierada, R. D. Yusen, and J. D. Cooper, ''Patients with emphysema: quantitative CT analysis before and after lung volume reduction surgery,'' Radiology **203**, 705–714 (1997).

[11]T. Pavlidis, *Algorithms for Graphics and Image Processing* (Springer-Verlag, Berlin, 1982).

[12]M. D. Hayhurst, D. C. Flenley, A. McLean, A. J. A. Wightman, W. MacNee, D. Wright, D. Lamb, and J. Best, ''Diagnosis of pulmonary emphysema by computerised tomography,'' Lancet **2**, 320–322 (1984).