

A texture analysis approach to quantify ventilation changes in hyperpolarised ^3He MRI of the rat lung in an asthma model

Frank Risse^{a*}, Jelena Pesic^a, Simon Young^b and Lars E. Olsson^{a,c}

In preclinical research, allergic asthma is investigated in rats sensitised with the antigen ovalbumin (OVA), followed by a challenge with aerosolised OVA to induce an inflammatory reaction of the lower airways. This causes diffuse, nonfocal ventilation defects that lead to heterogeneously distributed signal intensities in hyperpolarised (HP) ^3He MR images, which are difficult to assess directly by diagnostic grading or volumetry. Texture analysis can characterise these changes and does not require segmentation of the lung structures prior to the analysis. The aim of this work was to evaluate a texture analysis approach to quantify changes in lung ventilation in HP ^3He MRI of OVA-challenged rats. OVA-challenged animals were treated with two different compound doses to evaluate the sensitivity of the texture analysis. Four groups were investigated using HP ^3He MRI at 4.7 T: controls, vehicle-treated, and low- and high-dose budesonide-treated rats. In addition, broncho-alveolar lavage was performed and the eosinophil cell count was used as a biological reference marker. First-order texture, geometrical features and features based on second-order statistics using run-length and grey-level co-occurrence matrices were calculated. In addition, wavelet transforms were applied to compute first-order statistics on multiple scales. The texture analysis was able to show significant differences between the control and untreated vehicle groups as well as between the vehicle and treatment groups. This is in agreement with the findings of the eosinophil cell counts, which were used as a marker for the severity of inflammation. However, not all features used in the different texture analysis methods could differentiate between the treatment groups. In conclusion, texture analysis can be used to quantify changes in lung ventilation as measured with HP ^3He MRI after therapeutic intervention with budesonide. Copyright © 2011 John Wiley & Sons, Ltd.

Keywords: MRI; texture analysis; hyperpolarised; asthma; lung; rat; ovalbumin

INTRODUCTION

Asthma is the most common chronic respiratory disease and its prevalence is increasing. The disease is characterised by attacks of breathlessness and wheezing of varying severity, which are caused by smooth-muscle constriction of the airways and bronchial inflammation (1). Although animal models can only reproduce some aspects of human asthma, e.g. bronchial inflammation and airway remodelling (2,3), they can provide a link between *in vitro* experiments and human discovery and development trials of novel therapeutic agents. Sensitisation of rodents with the antigen ovalbumin (OVA), with a subsequent OVA challenge, is an example of such a model in allergic asthma research (3,4). OVA induces an inflammatory reaction of the lower airways (4,5), which causes impaired lung function.

In contrast with ^1H MRI, hyperpolarised (HP) ^3He MRI of the lungs provides direct imaging of the airspaces (6,7) and can be used to assess pulmonary ventilation (8,9) and respiratory dynamics (10). It is also a promising tool to study animal models of lung diseases (9,11). Recently, HP ^3He MRI has been used to investigate the effects of asthma on ventilation in small animals (11–13). Analysis of the acquired images in these studies was performed by visual inspection and assessment of the type (fissure defect, lobar defect, etc.), number and lesion area (11), calculating the number of ventilated voxels before and after allergen challenge (12), or calculating the fractional ventilation in each voxel (13). However, a challenge with inhaled OVA may result in diffuse, nonfocal ventilation defects, which lead

to heterogeneously distributed signal intensities in HP ^3He MR images. Therefore, it is difficult to assess changes in ventilation after therapeutic intervention directly by segmentation or counting. Even a grading of the images may not be possible if only small differences occur between animals. Furthermore, the additional use of images acquired before challenge is often not feasible in one imaging session as the maximum effect of the inflammation occurs after several hours (14). Therefore, two separate imaging sessions are necessary and the resulting images need to be registered. In addition, HP ^3He must be available at both imaging sessions, which is often not the case when no hyperpolarisation facility is available on site.

* Correspondence to: F. Risse, DECS Imaging&Antibodies, AstraZeneca R&D, SE-431 83 Mölndal, Sweden.
E-mail: frank.risse@astrazeneca.com

a F. Risse, J. Pesic, L. E. Olsson
DECS Imaging&Antibodies, AstraZeneca R&D, Mölndal, Sweden

b S. Young
Bioscience, AstraZeneca R&D, Charnwood, UK

c L. E. Olsson
Department of Radiation Physics, University of Gothenburg, Göteborg, Sweden

Abbreviations used: ANOVA, analysis of variance; BAL, broncho-alveolar lavage; Bud, budesonide; DWT, discrete wavelet transform; GLCM, grey-level co-occurrence matrix; HP, hyperpolarised; OVA, ovalbumin; p_{max} , maximum probability; Veh, vehicle.

Texture refers to properties describing the surface or structure of objects (15). Although no formal definition exists (16,17), the analysis of texture is of importance in pattern recognition and image processing (17). It has been used successfully to classify lung pathologies in computed tomography images (18–21), where structural changes cause changes in the signal distribution patterns of the images, and in HP ^3He MRI, where the use of the lacunarity provides a more sensitive measure of emphysema progression than does standard apparent diffusion coefficient analysis (22). Recently, texture analysis has been applied directly to HP ^3He MR images in asthmatic and nonasthmatic patients. It was shown that some of the calculated features had a higher diagnostic relevance than clinical standard spirometry (23).

Texture analysis should also be well suited to quantify changes in ventilation caused by drug treatment measured with HP ^3He MRI, because it can be assumed that the heterogeneity of ventilation, and thus the signal pattern, will change on treatment. Texture analysis can characterise these changes directly and does not require segmentation of the lung structures. Furthermore, it is not operator dependent and can be executed automatically. The aim of this work was to evaluate a texture analysis approach to quantify changes in lung ventilation in HP ^3He MRI of OVA-challenged rats. OVA-challenged animals were treated with two different compound doses to evaluate the sensitivity of the texture analysis.

METHODS

Animal preparation and image acquisition

The study was approved by the Gothenburg local ethical committee on animal experiments. Twenty-four male Brown-Norway rats (weight, 175–200 g) were divided into four groups containing six animals each (Table 1). On day 0, the animals were sensitised with a subcutaneous injection of 500 µg OVA absorbed onto 100 mg aluminium hydroxide in 0.4 mL deionised water that was administered to each animal at two distinct sites with approximately 0.2 mL per site. On day 14, the rats were challenged with aerosolised OVA for 15 min. The rats were placed in groups of up to eight in a polyvinyl chloride inhalation box during the challenge (internal dimensions, 55 × 35 × 20 cm³). Two jet nebulisers (Pari LC Jet Star, PariPharma, Munich, Germany) were filled with 5 mL of 10 mg/mL OVA in 0.9% saline or 0.9% saline alone. Compressed air at 6 L/min flowed through each nebuliser and the output of the nebulisers was passed into the box containing the rats.

Table 1. Study groups

Group	No. of rats	Sensitisation	Challenge	Treatment
Control	6	OVA	Saline	Veh
OVA/Veh	6	OVA	OVA	Veh
OVA/Bud 0.1	6	OVA	OVA	0.1 mg/kg Bud
OVA/Bud 1.0	6	OVA	OVA	1.0 mg/kg Bud

Bud, budesonide; OVA, ovalbumin; Veh, vehicle.

Vehicle (Veh) (saline solution, NaCl 9 mg/mL, Fresenius Kabi, Oslo, Norway) or budesonide (Bud) (Sigma-Aldrich, Stockholm, Sweden) was administered intratracheally using a metal cannula with a bulb-formed head. The rats were treated twice daily, one treatment in the morning and one in the afternoon with 6 h in between, starting 2 h before the OVA challenge, using a volume of 0.5 mL/kg with a Bud concentration of either 0.1 mg/kg (OVA/Bud 0.1) or 1 mg/kg (OVA/Bud 1.0).

The imaging experiments were performed 48 h after OVA challenge. The animals were anaesthetised with 90 mg/kg ketamine (Ketalar, Pfizer AB, Sollentuna, Sweden) and 10 mg/kg intraperitoneal xylazine (Rompun, Bayer AB, Göteborg, Sweden). A tracheal intubation was performed using an intravenous catheter (BD Venflon, Becton Dickinson, 20 GA, Infusion Therapy AB, Helsingborg, Sweden) as tracheal tube. In addition, 0.1 mg/kg intraperitoneal pancuronium (Pavulon, Organon, Teknika, Boxtel, the Netherlands) was administered as muscle relaxant. The animals were sacrificed immediately after the experiments using 0.5 mL pentobarbitone sodium (60 mg/mL, Apoteksbolaget, Solna, Sweden). Imaging was performed on a BioSpec 47/40 4.7-T MR scanner (Bruker BioSpin, Ettlingen, Germany) equipped with a BGA12 gradient insert (maximum strength, 400 mT/m, Ø = 120 mm) and a double-tuned ^1H (linear) and ^3He (quadrature) transmit/receive coil (Ø = 72 mm). The scanner was equipped with an animal bed with a fluid-based heating system driven by a thermostat-controlled pump (Haake DC 10, Thermo Electron Corporation, Philadelphia, PA, USA), which prevented the animals from cooling down during imaging. The HP ^3He images were acquired in three breath-holds using a three-dimensional fast low-angle shot pulse sequence with TE = 1.0 ms, TR = 2.4 ms, flip angle of 3.5°, bandwidth of 75 kHz, field of view of 50 × 30 × 36 mm³ and a matrix size of 96 × 96 × 24.

A small-animal ventilator system dedicated for HP ^3He was used during the experiment (24). The ventilator was connected to pressurised air and was controlled by a computer running Windows XP (Microsoft Inc., Redmond, WA, USA). Nitrogen gas and a vacuum pump for rinsing the system were also connected. During the animal experiment, HP ^3He was temporarily stored in a Tedlar bag located in a box inside the magnet. Under normal conditions, the animal was set to breathe air with a respiratory rate of 60–70/min at a tidal volume of 2–2.5 mL. For ^3He image acquisition, the ventilator was set to ^3He delivery and the images were acquired in segments in three different respiratory cycles of approximately 2 s, each with a 1.2-s breath-hold. Linear k -space acquisition was performed and the k -space centre was acquired during the middle of the second breath-hold. The acquisition of the images was preceded by one additional breathing cycle of ^3He gas in order to improve the signal-to-noise ratio. The peak inspiration pressure was kept constant at $P = 15$ mbar and the tidal volume was allowed to change, ensuring similarly inflated lungs independent of the degree of oedema.

A broncho-alveolar lavage (BAL) was performed after imaging to collect inflammatory cells. The number of eosinophils in the BAL fluid was counted and used as reference for the texture analysis.

Image preprocessing

All image processing methods were implemented in Interactive Data Language (IDL 6.3, ITT Visual Solutions, Boulder, CO, USA) and were run without user interaction.

A threshold needs to be applied to remove image background, which would otherwise be included in the computation of textural features. This was performed using the maximum entropy threshold method (25), which is an algorithm to determine the position of the maximum entropy in the image histogram in order to set an optimal threshold. It should be noted that ventilation defects in HP ^3He MRI have a very low signal-to-noise ratio and may also be removed when the threshold is applied.

The second preprocessing step applies only to the calculation of second-order features as described below. Both methods used rely on local grey-level similarities. However, the original 16-bit signed data format of the MR images has a grey-level range from 0 to 32 768 and therefore a very low probability of finding an adequate number of voxels with the same value. To reduce the number of grey levels, image binning $s = \text{floor}[s_{\text{max}}(i - i_{\text{min}})/(i_{\text{max}} - i_{\text{min}})]$ was used to reduce the grey-level range to 64 values. s is the scaled intensity, s_{max} is the maximum of the scaled image intensity (here $s_{\text{max}} = 64$), i is the image intensity, i_{min} is the minimum intensity and i_{max} is the maximum intensity.

Feature extraction

First- and second-order statistical textural features and geometrical features were calculated to quantify the ventilation defects. First-order features were derived directly from the images after the removal of background. Second-order features were calculated on the basis of co-occurrence or run-length matrices after image binning.

First-order statistics and geometrical features

First-order features represent the distribution of grey levels in the images, e.g. mean, standard deviation and entropy (18). A list of all first-order statistical features used and their definitions is given in Table 2. In addition, simple geometrical features, such as volume and surface area to volume ratio, were calculated. It can be expected that nonventilated regions, which are removed in the first preprocessing step, will lead to a reduced total ventilated lung volume and an increased lung surface as a result of holes and irregularities.

Run length

The use of run length or primitive length to describe image texture was introduced by Galloway (26). Starting from a voxel with a certain grey level s , a run will continue in the direction ϕ as

long as the voxels along this direction have the same grey level s , i.e. the run will stop as soon as a different grey level is reached. The detected image element is called a primitive, which can be described by its grey level, length and direction (15). Coarse textures are therefore characterised by long primitives and fine textures by shorter primitives.

For computational efficiency, each run was performed in the positive and negative direction at the same time, i.e. in the direction ϕ and in the direction of its negative supplementary angle $\phi - 180^\circ$. Voxels that already contributed to a primitive in direction ϕ_i were excluded from any other run in the same direction. Therefore, one particular voxel can only belong to one particular primitive per direction. All diagonal run lengths were multiplied by $\sqrt{2}$ as proposed by Galloway (26).

A matrix $R_\phi = [r_{s,l}]_{K,L}$, where K is the number of grey levels and L is the maximum run length, containing the number of primitives with length l and grey level s , was computed for each ϕ . The total number of runs in a three-dimensional ($M \times N \times Z$) image is then:

$$R = \sum_{\phi} R_{\phi} = \sum_{s=1}^K \sum_{l=1}^L r(s, l) \quad [1]$$

The run lengths were calculated in four directions, $\phi = [0^\circ, 45^\circ, 90^\circ, 135^\circ]$, and the maximum run length was restricted to $L = 16$ mm. Each run length $r(s, l)$ was sorted into the matrix R according to its physical length, and all directions were considered as being equivalent. A list with all features calculated on the basis of the run-length matrix R can be found in Table 3 (15,26). The features were normalised to the mean run length to make them scale independent.

Co-occurrence matrices

Grey-level co-occurrence matrices (GLCMs) are based on second-order statistics (17). They represent the spatial relationship of grey-level pairs in an image, i.e. the frequency of occurrence of pairs of grey levels with a certain distance between them along a certain direction (17). Formally, a GLCM of an $M \times N$ image can be described as a matrix of relative frequencies $P_{\phi,d}(s_1, s_2)$ (15)

$$P_{\phi,d}(s_1, s_2) = \frac{|[(i, j), (m, n)] \in D : i - m = 0, |j - n| = d, f(i, j) = s_1, f(m, n) = s_2|}{\sum_{s_1, s_2} |[(i, j), (m, n)] \in D : i - m = 0, |j - n| = d, f(i, j) = s_1, f(m, n) = s_2|} \quad [2]$$

where $|\dots|$ refers to cardinality, ϕ to the direction or angle in degrees, d to the distance and $D = (M \times N) \times (M \times N)$.

Table 2. First-order statistical and geometrical features

Number of voxels N	Number of voxels in the ventilated lung
Lung volume V	$V = N \times \text{voxel size}$
Mean signal intensity (or grey level) I	$I = \frac{1}{N} \sum_{k=0}^{N-1} i_k$
Coefficient of variation CV	$CV = \frac{SD}{I}$
Entropy H	$H = -\sum_{k=0}^{N-1} p(i_k) \log_2 p(i_k)$
	$p(i_k)$ is the probability density of all grey levels i in the image
Uniformity U	$U = \sum_{k=0}^{N-1} p(i_k)^2$
Surface S/V	Surface area of the ventilated lungs normalised to V

Table 3. Run-length features

Short primitives emphasis	$SPE = \frac{1}{R} \sum_{s=1}^K \sum_{l=1}^L \frac{r(s, l)}{l^2}$
Long primitives emphasis	$LPE = \frac{1}{R} \sum_{s=1}^K \sum_{l=1}^L r(s, l) l^2$
Grey-level uniformity	$GLU = \frac{1}{R} \sum_{s=1}^K \left[\sum_{l=1}^L r(s, l) \right]^2$
Primitive length uniformity	$PLU = \frac{1}{R} \sum_{l=1}^L \left[\sum_{s=1}^K r(s, l) \right]^2$
Primitive percentage	$PP = \frac{R}{\sum_{l=1}^L \sum_{s=1}^K} l \cdot r(s, l) = \frac{R}{MNZ}$

GLCMs can be computed in two or three dimensions. Here, only two dimensions were used because of the distinctly lower spatial resolution in the third dimension. Therefore, only four directions $\phi = [0^\circ, 45^\circ, 90^\circ, 135^\circ]$ and 12 distances ranging from 0.23 to 30 mm were used to compute the GLCM. It should be noted that, in the literature, sometimes eight directions are denoted when two-dimensional GLCMs are computed. This is because the 'negative' direction along ϕ is counted as an extra direction. However, because of the matrix symmetry of $P_{\phi, d}(s_1, s_2)$, both directions will cover the same information and are therefore not counted twice here. Each entry in the GLCM was normalised to the total number of paired occurrences. Textural features were derived from the GLCM according to refs. (15,16). A complete list of the derived features is given in Table 4.

Multi-scale first-order statistics (wavelets)

Wavelets can be used to represent functions by dilations and translations of a mother wavelet, similar to the Fourier transform, where cosine and sine functions are used. The main difference is that wavelets are 'local' in the space (image) domain and frequency domain, i.e. a change in the frequency domain leads to

a localised change in the image domain, whereas a change in the frequency domain would affect the entire image domain when a Fourier transform is used. Wavelet transform is mainly applied in image compression, but has also been employed successfully in texture analysis (27). In contrast with the run-length and GLCM features, which are restricted to a single scale and a relatively small neighbourhood, wavelet features work on multiple scales and can assess texture patterns in larger neighbourhoods (28).

An efficient way of computing the wavelet transform is the discrete wavelet transform (DWT), where dilations and translations are based on the power of two (29). In the standard approach, the image is decomposed in sub-bands (or subimages) by filtering and downsampling. The resulting subimages contain information on specific levels (scales) and orientations. Textural features can be calculated directly on the basis of the resulting subimages and will represent image properties at different scales and in different directions.

In the DWT described above, the next level is always computed from the so-called approximation subimage, which corresponds to coarse level coefficients, and the features of all subimages are analysed, leading to a large number of features. Chang and Kuo (27) introduced texture analysis with tree-structured DWT, in which only the subimage with the highest energy content E_{sub}

$$E_{\text{sub}} = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} i(x, y) \quad [3]$$

is used for further analysis, where $M \times N$ is the image matrix size and i is the signal intensity. This reduces the number of features considerably. The number of levels here was set to four to avoid subimages that were too small and unstable results. The mother wavelet used for the DWT was a second-order Debauchies wavelet, and zero-filling was applied to achieve image dimensions the size of the next greatest power of two. The DWT was performed directly on the original images, i.e. no binning was carried out.

Table 4. Grey-level co-occurrence matrix (GLCM) features

Energy	$E = \sum_{s_1, s_2} P_{\phi, d}^2(s_1, s_2)$	
Entropy	$H_{\text{co}} = \sum_{s_1, s_2} P_{\phi, d}(s_1, s_2) \log(P_{\phi, d}(s_1, s_2))$	
Maximum probability	$p_{\text{max}} = \max_{s_1, s_2} P_{\phi, d}(s_1, s_2)$	
Contrast	$C = \sum_{s_1, s_2} s_1 - s_2 ^\kappa P_{\phi, d}^\lambda(s_1, s_2)$ $\kappa = 2, \lambda = 1$	
Inverse difference moment	$\text{IDM} = \sum_{s_1, s_2, i \neq s_2} \frac{P_{\phi, d}^\lambda(s_1, s_2)}{ s_1 - s_2 ^\kappa}$ $\kappa = 2, \lambda = 1$	
Correlation	$r = \frac{\sum_{s_1, s_2} [(s_1 s_2) P_{\phi, d}(s_1, s_2)] - \mu_x \mu_y}{\sigma_x \sigma_y}$ $\mu_x = \sum_{s_1} s_1 \sum_{s_2} P_{\phi, d}(s_1, s_2)$ $\mu_y = \sum_{s_2} s_2 \sum_{s_1} P_{\phi, d}(s_1, s_2)$	$\sigma_x = \sum_{s_1} (s_1 - \mu_x)^2 \sum_{s_2} P_{\phi, d}(s_1, s_2)$ $\sigma_y = \sum_{s_2} (s_2 - \mu_y)^2 \sum_{s_1} P_{\phi, d}(s_1, s_2)$

Statistics

All statistical calculations, including the results of all animals used, were performed with Prism 4 for Windows (GraphPad Software Inc., La Jolla, CA, USA) using one-way analysis of variance (ANOVA) with Dunnett's multiple comparison test against the Veh group as post-test. The different textural features were also correlated with the number of eosinophils from the BAL procedure using Pearson's linear correlation. $p < 0.05$ was defined as significant.

RESULTS

Figure 1 shows typical example slices from one control and three OVA-challenged animals. The control animals showed a relatively homogeneous distribution of ^3He in the lungs, where the low signal in the central lung regions mostly corresponds to major blood vessels. However, some regions may also be caused by inflammation unrelated to OVA as Brown-Norway rats are prone to underlying lung inflammation. In contrast with the controls, the OVA-challenged and Veh-treated (OVA/Veh) animals showed a heterogeneously distributed signal and focal regions with a very low signal. This was also seen, although to a lesser extent, in the OVA/Bud 0.1 group. The high-dose Bud-treated (OVA/Bud 1.0) group showed clearly delineated lungs with a homogeneous signal distribution.

The number of eosinophils in the OVA/Veh group was increased significantly relative to that in the control group (24-fold, $p < 0.001$). The number of eosinophils in the Bud-treated groups was reduced two-fold ($p < 0.05$) for the low-dose and 14-fold ($p < 0.001$) for the high-dose groups relative to the Veh-treated group (Fig. 2).

Texture analysis was performed to quantify the differences between the groups. The calculation of all textural features in a three-dimensional dataset of one animal took approximately 1 min on a conventional personal computer (Intel P4, 3.6 GHz, 2 GB RAM, Windows XP). All geometrical and first-order statistical features, except the mean intensity, showed

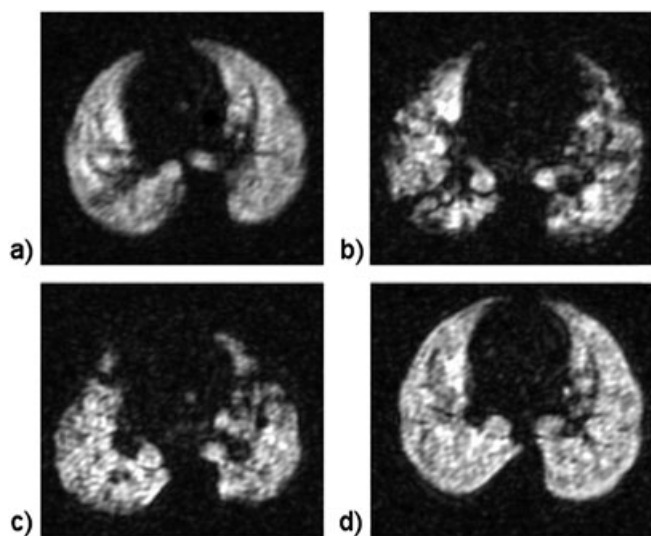


Figure 1. Example images of the lung ventilation as acquired with hyperpolarised (HP) ^3He MRI in the four study groups: (a) control; (b) vehicle-treated group; (c) low-dose budesonide group; (d) high-dose budesonide group.

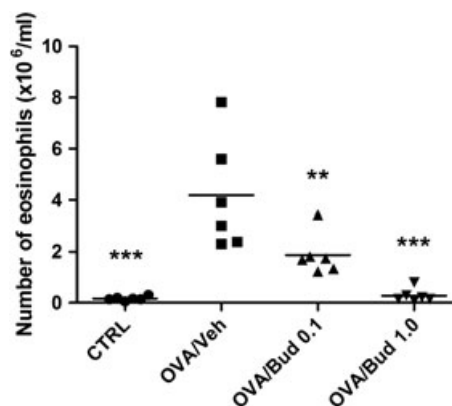


Figure 2. Comparison of the groups using the number of eosinophils collected with broncho-alveolar lavage. The control (CTRL) and both treatment groups are significantly different from the vehicle-treated group [one-way analysis of variance (ANOVA) with Dunnett's multiple comparison test: *** $p < 0.001$, ** $p < 0.01$]. Bud, budesonide; OVA, ovalbumin; Veh, vehicle.

significant differences between the OVA/Veh group and the control and OVA/Bud 1.0 groups. Only the surface area/volume ratio showed a significant difference between the OVA/Veh and the OVA/Bud 0.1 group, which was also found in the cell counts (Fig. 3).

Figure 4 shows the comparison between the groups using the results from short primitive emphasis and the primitive percentage features as examples for the run-length methods, as all features revealed the same significance levels. All run-length features found highly significant differences between the OVA/Veh group and the control and OVA/Bud 1.0 groups. However, no significant differences were seen when the low-dose OVA/Bud 0.1 group was compared with the OVA/Veh group.

The GLCM calculation results in 288 features that describe the texture of the images (four directions $\times$ 12 distances $\times$ six features), too many for a comparison of the study groups. The feature information was therefore summarised in two steps: (i) the feature mean value over the four directions was calculated for each distance; this results in a distribution of the features over the distances; (ii) the resulting distribution was then characterised by its mean and skewness to determine whether shorter or longer distances were dominant. Figure 5 shows the results of the group comparison for the energy feature. All GLCM features are summarised in Fig. 6 by showing the mean differences between the OVA/Veh group and the other groups for each feature, including their significance levels. Only the contrast feature did not show significant differences between the OVA/Veh group and all other groups when characterised by the mean of the distribution. The other GLCM features revealed differences at least for the control and OVA/Bud 1.0 groups. Higher significance levels were generally found with the GLCM when the feature distribution was characterised by its skewness. Even the contrast feature showed significant differences for all groups, whereas the maximum probability $p_{\max}$ did not reveal significant differences.

The DWT-based features showed very similar results for all evaluated sub-bands, i.e. those with the highest energy content. They all revealed significant differences between the OVA/Veh group and the other groups. An example is shown in Fig. 7.

The linear correlation between the eosinophil counts as reference and the textural features across the measurements of all

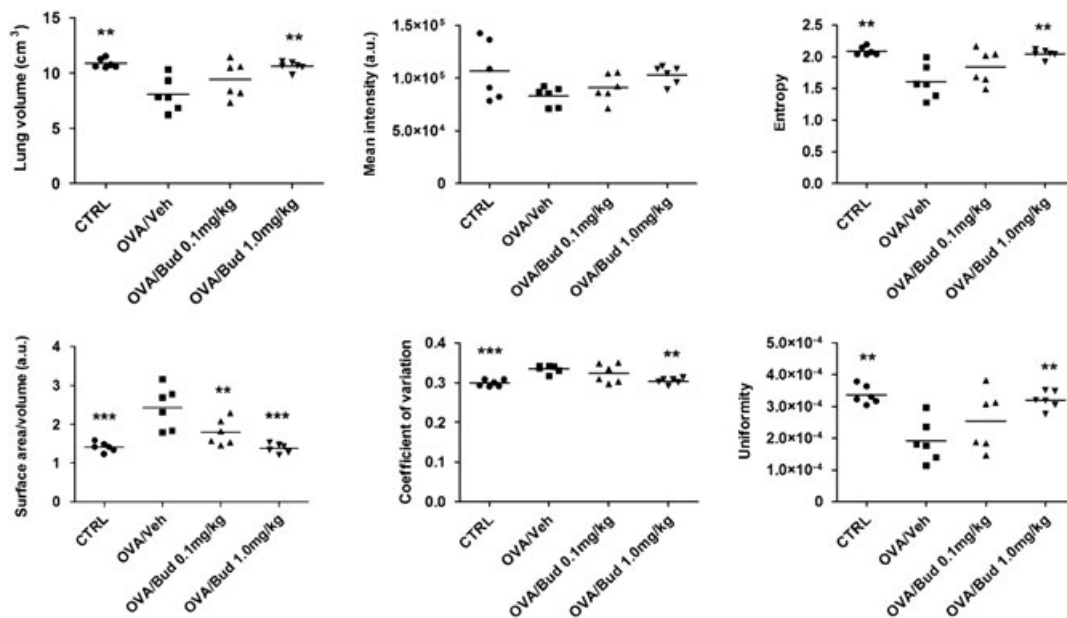


Figure 3. Comparison of the control (CTRL) and budesonide-treated groups with the vehicle-treated group using geometrical and first-order statistical features. Top row (left to right): ventilated lung volume, mean signal intensity and entropy. Bottom row (left to right): surface area normalised to the ventilated lung volume, coefficient of variation and uniformity. Statistical comparisons are with the OVA/Veh group using one-way analysis of variance (ANOVA) with Dunett's multiple comparison test: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Bud, budesonide; OVA, ovalbumin; Veh, vehicle.

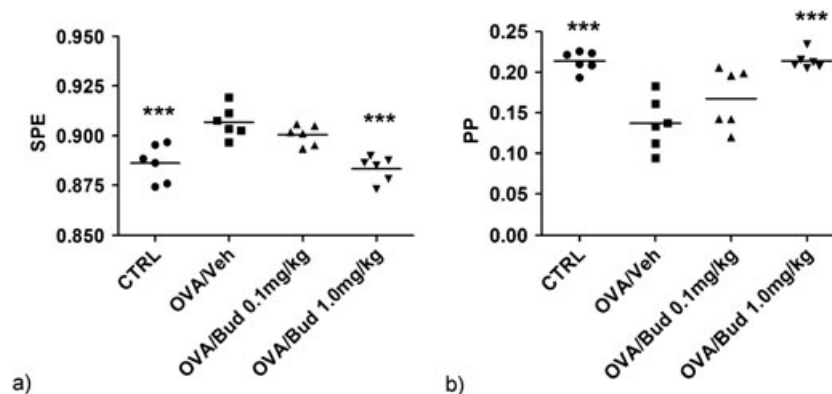


Figure 4. Comparison of the control (CTRL) and budesonide-treated groups with the vehicle-treated group using run-length features: (a) short primitive emphasis (SPE); (b) primitive percentage (PP). Statistical comparisons are with the OVA/Veh group using one-way analysis of variance (ANOVA) with Dunett's multiple comparison test: *** $p < 0.001$. Bud, budesonide; OVA, ovalbumin; Veh, vehicle.

animals, including the control group, resulted in correlations of $r = 0.82$ ($p < 0.0001$) for the surface area/volume ratio and $r = -0.76$ ($p < 0.0001$) for the lung volume. The correlation coefficients of the first-order features were in an absolute range of $|r| = 0.74$ – 0.79 ($p < 0.0001$), where the entropy and uniformity had a negative correlation. Only the signal mean showed a weak negative correlation of $r = -0.46$ ($p < 0.05$), which was outside this range. The run-length features revealed a negative correlation between $r = -0.70$ and $r = -0.79$ ($p < 0.001$), with the exception of the short primitive emphasis with $r = 0.73$ ($p < 0.0001$). Most of the GLCM features showed moderate to good correlation in the range of $|r| = 0.65$ – 0.80 ($p < 0.001$). However, no correlation between the eosinophil cell count and the textural features was found for two features: the maximum probability feature, when characterised by the skewness of its feature distribution, and the contrast feature, when characterised by the mean of its feature distribution.

Overall, the DWT-based features showed a moderate correlation of $r = 0.66$ ($p < 0.001$). Two examples of the correlation between features and the number of eosinophils are shown in Fig. 8.

DISCUSSION

This study demonstrates the feasibility of assessing changes in lung ventilation in allergic Brown-Norway rats after treatment with a corticosteroid using texture analysis of HP ^3He MR images. This method allows the quantification of ventilation images with nonfocal ventilation defects caused by an inhaled antigen, which is not possible with segmentation techniques and very difficult when image grading is used. Furthermore, it is time efficient and requires no user interaction, and should therefore lead to a higher reproducibility than manual or semi-automatic segmentation, as used in earlier studies

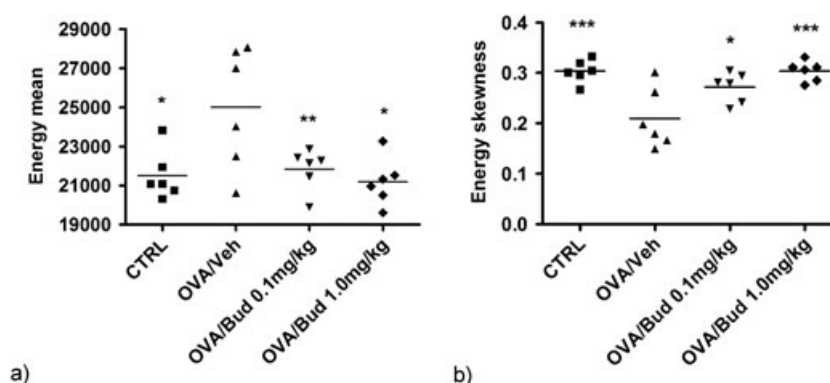


Figure 5. Comparison of the control (CTRL) and budesonide-treated groups with the vehicle-treated group using grey-level co-occurrence matrix (GLCM) features: (a) mean of the energy distribution of 12 distances; (b) skewness of the same energy distribution. Statistical comparisons are with the OVA/Veh group using one-way analysis of variance (ANOVA) with Dunett's multiple comparison test: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Bud, budesonide; OVA, ovalbumin; Veh, vehicle.

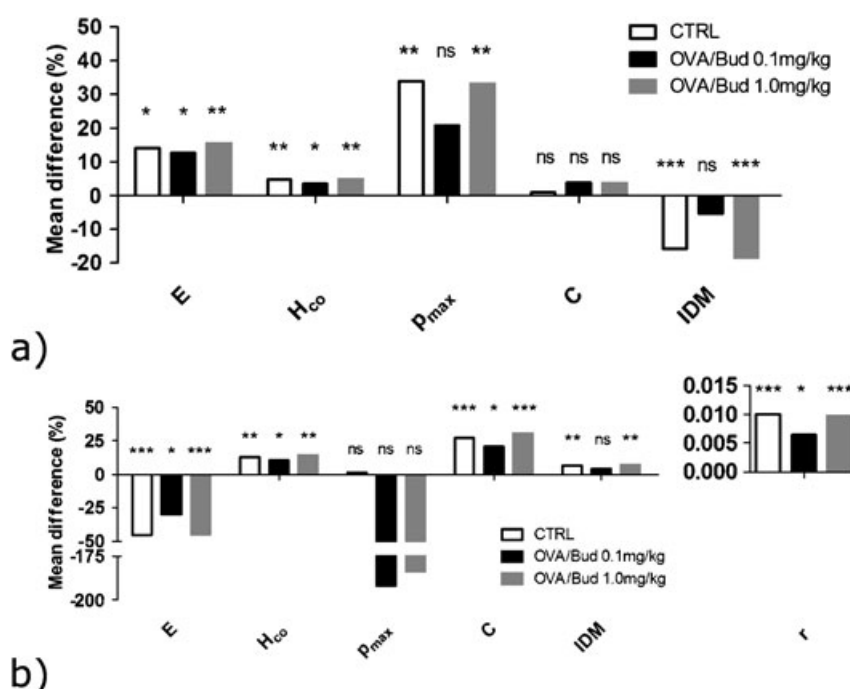


Figure 6. Mean differences between the vehicle (OVA/Veh) groups and the control (CTRL) and budesonide-treated groups calculated as: (a) the mean of the feature distribution over 12 distances; (b) skewness of the feature distribution (see Table 4 for an explanation of the abbreviations). The mean difference in the correlation r between the groups was zero for the mean of the feature distribution. Note the intercept and the different scale for the correlation r in (b). Statistical comparisons are with the OVA/Veh group using one-way analysis of variance (ANOVA) with Dunett's multiple comparison test: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Bud, budesonide; OVA, ovalbumin; Veh, vehicle.

with focal, clearly identifiable defects in a lipopolysaccharide model (9,30).

Nonfocal or diffuse ventilation defects were found in our study, which have not been described previously in the allergic Brown-Norway rat asthma model. In contrast, relatively clearly defined lesions have been described in previous studies using OVA for sensitisation and challenge in rats (31,32). The major difference is the use of T_1 -weighted ^1H MRI in these studies. Therefore, changes in ventilation are not detected directly, but rather by the appearance of oedema in the lungs. Furthermore, the administration of the OVA challenge was different, which may also have led to a different appearance of the inflammation. Beckmann *et al.* (31) used an OVA instillation, which may account

for the more focal lesions. Ble *et al.* (32) administered the OVA challenge intranasally after an intraperitoneal sensitisation in mice, which is also more similar to an instillation.

Four different feature classes were used to analyse the images. Most of the features used revealed significant differences between the Veh, i.e. untreated, group and both the control and Bud-treated groups, which is in agreement with the number of eosinophils recovered by the BAL fluid as reference measurement. However, relative differences between the Veh and other groups are not directly comparable between the cell counts and the image features. The number of eosinophils represents a surrogate measure for the severity of inflammation on a cellular level, whereas the image features characterise different image

properties or changed ventilation on a more macroscopic level. This may also explain the variations in the performance of the texture analysis methods and their features in identifying differences between the groups.

The geometrical features, which are not, strictly speaking, textural features, were assumed to be a simple measure of the appearance of the lungs, i.e. size, shape regularity and number of nonventilated regions. The surface area/lung volume feature, which is related to the counting of ventilation defects, revealed significant differences even between the Veh (OVA/Veh) and low-dose Bud (OVA/Bud 0.1) group on the same significance levels as the cell counts. The feature represents the total surface area, including the inner surfaces, normalised to the ventilated lung volume. It can be assumed that a healthy lung has a lower surface/volume ratio (like a sphere) than a lung with ventilation defects, which better resembles a sponge. No significant difference in lung volume was shown between the OVA/Veh group and the OVA/Bud 0.1 group. This is probably because plasma leakage was not entirely inhibited at the low steroid dose, and therefore some animals showed reduced ventilated lung volumes similar to the Veh group. Furthermore, both groups

showed relatively high standard deviations, making the lung volume less sensitive to changes between the groups.

In general, the first-order statistics features showed significant differences only between the OVA/Veh and the control and high-dose Bud (OVA/Bud 1.0) groups. The mean signal intensity revealed no significant difference between any of the groups. All of these features are based on the intensity histogram rather than directly on image structures. The coefficient of variation, entropy and uniformity are related to the homogeneity of the image intensity, and thus also to the homogeneity of lung ventilation, which is expected to decrease with increasing severity of inflammation in the OVA model. The coefficient of variation has already been used successfully to describe heterogeneously distributed signals in pulmonary perfusion MRI (33).

The run-length method describes directly the structural properties of the ventilation images and is related to the extent of the ventilation defects. Healthy lungs should show relatively homogeneously distributed ventilation, and therefore longer primitives, than lungs with ventilation defects, which should show a larger short primitive emphasis. The run-length features showed significant differences between the OVA/Veh and control groups, and also between the OVA/Veh and OVA/Bud 1.0 groups. However, no difference was found between the OVA/Veh and OVA/Bud 0.1 groups. This lack of sensitivity may be caused by several effects. It could be caused by the reduction in the number of grey levels or signal intensities performed with image binning. Subtle changes in the image texture may be removed by reducing the grey levels to 64. A larger number of grey levels could lead to an improved sensitivity, but at the cost of computational efficiency. Split-and-merge algorithms (18,34), where voxels are merged on the basis of their image intensity and spatial relationship, may provide a better alternative to the binning approach used. Another reason could be that a decrease in the severity of inflammation does not translate directly to improved ventilation, which is thought to increase the homogeneity of the signal distribution. In this study, the three-dimensional datasets had a reduced spatial resolution in the z-direction. Isotropic three-dimensional images could improve the performance of the run-length method by exploiting the whole image information, as performed successfully in a recent study classifying emphysema on lung computed tomography images (18).

The GLCM method revealed inconsistent results for the evaluated features, e.g. the maximum probability showed significance when expressed as the mean over all distances, but not when expressed as the skewness. However, in contrast with the run-length features, some of the GLCM features found significant

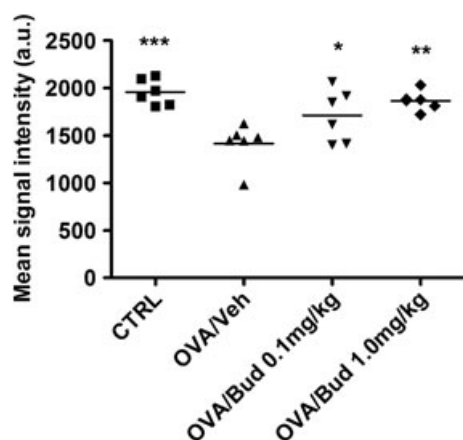


Figure 7. Comparison of the control (CTRL) and budesonide-treated groups with the vehicle-treated group using the mean signal values of the sub-image with the highest information content on the first level of the discrete wavelet transform (DWT). Statistical comparisons are with the OVA/Veh group using one-way analysis of variance (ANOVA) with Dunnett's multiple comparison test: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Bud, budesonide; OVA, ovalbumin; Veh, vehicle.

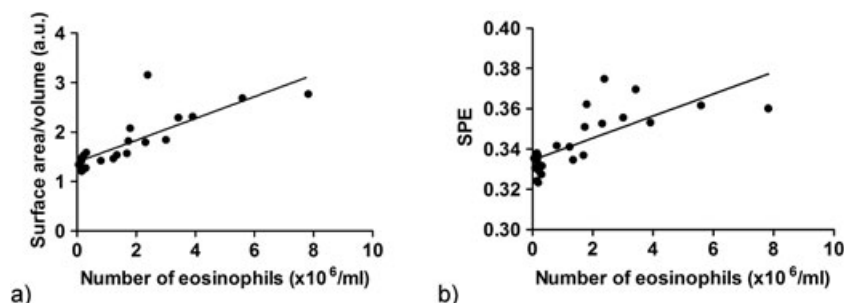


Figure 8. Linear correlation between the surface area/volume ratio (a) and the short primitive emphasis (SPE) (b) and the number of eosinophils. The correlation coefficients are $r = 0.82$ ($p < 0.0001$) (a) and $r = 0.68$ ($p = 0.0002$) (b).

differences not only between the OVA/Veh and control and OVA/Bud 1.0 groups, but also between the OVA/Veh and OVA/Bud 0.1 groups. However, other features did not reveal any significant differences between the groups, highlighting the different properties of the features. Gotlieb and Kreyszig (17) grouped the GLCM features as: (i) features expressing visual textural characteristics – energy, contrast and correlation; (ii) features based on statistics – maximum probability and inverse difference moment (or homogeneity); and (iii) features based on information theory – entropy. Features in group (i) are related directly to the visual appearance of the image and thus also to the changes in lung ventilation, whereas the other two groups are based on more theoretical measures: for example, a large inverse difference moment is related to a homogeneous ventilation, whereas the entropy will increase with the heterogeneity of the ventilation. However, the GLCM features do not describe exactly the same characteristics as the first-order statistical features with the same names, as they are based on second-order statistics considering the spatial relationship between the voxels. In the example of the maximum probability feature $p_{\max}$, it describes the probability for a co-occurrence of a grey level at a certain distance, i.e. it describes how coarse is the texture. Different distances were investigated herein and the probabilities for these distances are summarised using their distribution. It is very unlikely that the shape (skewness) of the distribution will change with the change in texture, but the mean value should change when the coarseness of the texture changes. Therefore, only the mean value reveals significant differences in this example.

The GLCM is implemented in two dimensions only and a three-dimensional implementation did not show better results, in contrast with Xu *et al.* (18). However, their three-dimensional GLCM was used to classify multi-detector computed tomography lung images with high spatial resolution in all three directions. Furthermore, the GLCM and its features are computed in four directions in the image plane, but only the sum matrix of all directions was used in the evaluation in order to reduce the number of features. The distances were also not optimised for the underlying texture. Optimised distances with higher weights on certain distance ranges could therefore lead to improved results (18). Moreover, the features from the different distances were summarised and expressed as the mean and skewness of the resulting distribution, which may lead to a poorer performance and make it more difficult to interpret the GLCM features. Thus, the GLCM method can provide additional information or support results from other features, but there is, in contrast with the other methods, no straightforward relationship with the underlying biological changes.

The results from both feature classes using second-order statistics, namely run length and GLCM, may be affected by the threshold applied to the images in order to remove the image background. Severe ventilation defects appear in the same grey-level range as background. Therefore, they are also treated as background and not included in the analysis. However, this applies more to the run-length than to the GLCM features. No run length was measured in these severe defect areas, whereas the GLCM takes background voxels indirectly into account as it cannot find a corresponding voxel in background areas. The texture analysis using run-length and GLCM features could be improved by defining the lung borders, e.g. by combining ^1H and ^3He information, if available, or the segmentation methods used by Tustison *et al.* (23), and applying the run-length and GLCM methods only in the defined lung regions but including all

grey levels. However, an automatic segmentation of the lungs is not trivial as ventilation defects often occur at the lung borders, which makes the correct definition of the lung outlines challenging.

The wavelet analysis works directly on the original images and on different levels, which should be beneficial as texture is strongly scale dependent (15,27). The mean value of the selected sub-images revealed significant differences for all groups on all levels used. Further studies with larger numbers of animals per group are needed to evaluate whether all levels are required. The DWT analysis was performed using a tree-structured DWT approach proposed by Chang and Kuo (27). Only the sub-image with the highest energy content is chosen for the image statistics on each level as this sub-image should contain the most relevant information. However, it should be investigated in detail whether the conventional DWT could deliver better results by avoiding the selection, despite the larger number of features to be evaluated. Further improvement may be feasible by optimising the choice of the wavelet. Only the second-order Debauchies wavelet was used here.

Considering only the textural features that revealed a significant difference between the groups, moderate to good correlations between these features and the eosinophil cell count, which is a measure of the severity of inflammation, were found. The textural features should therefore reflect the severity of the disease to a certain extent. However, the relationship between the changes on a cellular level, measured by the cell count, and the more macroscopic ventilation changes is not likely to be linear, as can be seen in Fig. 8b. A small number of inflammatory cells probably have no effect on the ventilation, whereas, in the upper range, the ventilation can only be impaired to a certain level even if the number of eosinophils is very high. This can be seen in Fig. 8b, where cell counts above $5 \times 10^6/\text{mL}$ do not lead to increased short primitive emphasis, i.e. smaller structures. In addition, oedema could cause ventilation defects in regions not directly affected by the inflammation. There may also be a delay between the inflammation severity, as measured by the cell count, and its manifestation in the acquired images, thereby impairing the relationship between textural features and cell counts. Overall, a linear relationship can only be expected within certain limits, but more experiments are necessary to determine this range. However, it can be demonstrated that the textural features and the number of inflammatory cells are not completely unrelated.

In addition to the methods used here, other textural features have been proposed in the literature and successfully applied in the analysis of lung images, e.g. lacunarity, a fractal feature, used on apparent diffusion coefficient maps (22). Future investigations are necessary to evaluate these features for use on HP ^3He ventilation images. However, a large number of textural features were employed here, and it is very likely that redundancies between the computed features occurred. It is therefore arguable whether all features are necessary to assess differences in the dose–response between groups. A set of features could be selected on the basis of the different image characteristics they describe. In the first-order statistical and geometrical group, S/V showed the best performance and is easy to interpret. In addition, a measure of the homogeneity of the intensity distribution, i.e. the coefficient of variation, entropy or energy, could be chosen, as these are closely related to each other. The same could be performed for the other texture analysis methods by selecting features that are based on different characteristics to avoid

possible redundancies. However, additional studies with a larger number of animals and different disease models are required to optimise the choice of textural features further. A better approach would be the evaluation of the features against a standard read-out, e.g. against the clinical diagnosis, using mutual information, as performed by Tustison *et al.* (23), which was not feasible because of the diffuse ventilation defects in our study. Tustison *et al.* (23) also presented a ranking of the features used in the same study, which were similar to the features employed here, demonstrating the relevance of the run-length and GLCM methods for the analysis of clinical asthma data acquired by HP ^3He MRI. However, in contrast with the analysis presented here, regional masks were generated and the resulting regions were analysed separately.

An advantage of the texture analysis presented here is that no baseline measurement is necessary to quantify the changes in ventilation, which simplifies the experimental procedure and saves expensive HP ^3He . In addition, only a single dataset is needed, thus avoiding the registration of datasets at different time points. Moreover, this facilitates the quantification of ventilation changes in longitudinal studies to assess disease progression, as it allows comparison of the texture analysis results directly. An example of such a longitudinal investigation is the HP ^3He MRI study in lipopolysaccharide-challenged mice performed by Olsson *et al.* (9).

The implemented texture analysis methods could also be used in the segmentation of lesions in the lung or other organs in MRI based on texture classification when more focal lesions occur. A feature space could be created from the calculated features and k nearest-neighbour, Bayesian or other classifiers could be applied for feature classification. Another potential application is the analysis of the heterogeneity of lung perfusion maps in contrast-enhanced MRI or oxygen-enhanced MRI, especially if global parameters are not changed as a result of compensation effects.

CONCLUSIONS

Texture analysis can be used to quantify changes in lung ventilation measured by HP ^3He MRI after therapeutic intervention with budesonide. This method enables the analysis of ventilation images with diffuse ventilation defects that cannot be segmented or counted directly. In addition, the texture analysis presented could also be applied to images showing focal ventilation defects, particularly when focal and diffuse defects occur together. Moreover, it allows HP ^3He MR images to be analysed in an efficient manner with no user interaction during processing.

REFERENCES

- Fanta CH. Asthma. *N. Engl. J. Med.* 2009; 360: 1002–1014.
- Lloyd CM, Robinson DS. Allergen-induced airway remodelling. *Eur. Respir. J.* 2007; 29: 1020–1032.
- Fuchs B, Braun A. Improved mouse models of allergy and allergic asthma – chances beyond ovalbumin. *Curr. Drug Targets* 2008; 9: 495–502.
- Tigani B, Cannet C, Karmouty-Quintana H, Ble FX, Zurbrugg S, Schaeublin E, Fozard JR, Beckmann N. Lung inflammation and vascular remodeling after repeated allergen challenge detected noninvasively by MRI. *Am. J. Physiol. Lung Cell. Mol. Physiol.* 2007; 292: L644–L653.
- Hylkema MN, Hoekstra MO, Luinge M, Timens W. The strength of the OVA-induced airway inflammation in rats is strain dependent. *Clin. Exp. Immunol.* 2002; 129: 390–396.
- Bachert P, Schad LR, Bock M, Knopp MV, Ebert M, Grossmann T, Heilw, Hofmann D, Surkau R, Otten EW. Nuclear magnetic resonance imaging of airways in humans with use of hyperpolarized ^3He . *Magn. Reson. Med.* 1996; 36: 192–196.
- Middleton H, Black RD, Saam B, Cates GD, Cofer GP, Guenther R, Happer W, Hedlund LW, Johnson GA, Juvan K. MR imaging with hyperpolarized ^3He gas. *Magn. Reson. Med.* 1995; 33: 271–275.
- Woodhouse N, Wild JM, Paley MN, Fichelle S, Said Z, Swift AJ, van Beek EJ. Combined helium-3/proton magnetic resonance imaging measurement of ventilated lung volumes in smokers compared to never-smokers. *J. Magn. Reson. Imaging* 2005; 21: 365–369.
- Olsson LE, Smailagic A, Onnervik PO, Hockings PD. (1)H and hyperpolarized (3)He MR imaging of mouse with LPS-induced inflammation. *J. Magn. Reson. Imaging* 2009; 29: 977–981.
- Deninger AJ, Mansson S, Petersson JS, Pettersson G, Magnusson P, Svensson J, Fridlund B, Hansson G, Erjefeldt I, Wollmer P, Golman K. Quantitative measurement of regional lung ventilation using ^3He MRI. *Magn. Reson. Med.* 2002; 48: 223–232.
- Thomas AC, Potts EN, Chen BT, Slipetz DM, Foster WM, Driehuys B. A robust protocol for regional evaluation of methacholine challenge in mouse models of allergic asthma using hyperpolarized ^3He MRI. *NMR Biomed.* 2009; 22: 502–515.
- Driehuys B, Walker J, Pollaro J, Cofer GP, Mistry N, Schwartz D, Johnson GA. ^3He MRI in mouse models of asthma. *Magn. Reson. Med.* 2007; 58: 893–900.
- Haczku A, Emami K, Fischer MC, Kadlecsek S, Ishii M, Panettieri RA, Rizi RR. Hyperpolarized ^3He MRI in asthma measurements of regional ventilation following allergic sensitization and challenge in mice – preliminary results. *Acad. Radiol.* 2005; 12: 1362–1370.
- Beckmann N, Tigani B, Sugar R, Jackson AD, Jones G, Mazzoni L, Fozard JR. Noninvasive detection of endotoxin-induced mucus hypersecretion in rat lung by MRI. *Am. J. Physiol. Lung Cell. Mol. Physiol.* 2002; 283: L22–L30.
- Sonka M, Hlavac V, Boyle R. *Image Processing, Analysis, and Machine Vision*. Thomson Engineering: Toronto, ON; 2007.
- Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. *IEEE Trans. Systems Man Cybernetics* 1973; 3: 610–621.
- Gotlieb CC, Kreyszig HE. Texture descriptors based on co-occurrence matrices. *Comput. Vis. Graph. Image Process.* 1990; 51: 70–86.
- Xu Y, Sonka M, McLennan G, Guo J, Hoffman EA. MDCT-based 3-D texture classification of emphysema and early smoking related lung pathologies. *IEEE Trans. Med. Imaging* 2006; 25: 464–475.
- Boehm HF, Fink C, Attenberger U, Becker C, Behr J, Reiser M. Automated classification of normal and pathologic pulmonary tissue by topological texture features extracted from multi-detector CT in 3D. *Eur. Radiol.* 2008; 18: 2745–2755.
- Al-Kadi OS, Watson D. Texture analysis of aggressive and nonaggressive lung tumor CE CT images. *IEEE Trans. Biomed. Eng.* 2008; 55: 1822–1830.
- Park YS, Seo JB, Kim N, Chae EJ, Oh YM, Lee SD, Lee Y, Kang SH. Texture-based quantification of pulmonary emphysema on high-resolution computed tomography: comparison with density-based quantification and correlation with pulmonary function test. *Invest. Radiol.* 2008; 43: 395–402.
- Diaz S, Casselbrant I, Piitulainen E, Magnusson P, Peterson B, Pickering E, Tuthill T, Ekberg O, Akeson P. Progression of emphysema in a 12-month hyperpolarized ^3He -MRI study: lacunarity analysis provided a more sensitive measure than standard ADC analysis. *Acad. Radiol.* 2009; 16: 700–707.
- Tustison NJ, Altes TA, Song G, de Lange EE, Mugler JP, III, JC Gee. Feature analysis of hyperpolarized helium-3 pulmonary MRI: a study of asthmatics versus nonasthmatics. *Magn. Reson. Med.* 2010; 63: 1448–1455.
- Perez De Alejo R, Ruiz-Cabello J, Palmira V, Ignacio R, Perez-Sanchez JM, Peces-Barba G, Cortijo M. A fully MR-compatible animal ventilator for special-gas mixing applications. *Concepts Magn. Reson.* 2005; 26B: 93–103.
- Sahoo PK, Soltani S, Wong KC, Chen YC. A survey of thresholding techniques. *Comput. Vis. Graph. Image Process.* 1988; 41: 233–260.
- Galloway MM. Texture classification using gray level run length. *Comput. Graph. Image Process.* 1975; 4: 172–179.
- Chang T, Kuo CJ. Texture analysis and classification with tree-structured wavelet transform. *IEEE Trans. Image Process.* 1993; 2: 429–441.

28. Arivazhagan S, Ganesan L. Texture classification using wavelet transform. *Pattern Recog. Lett.* 2003; 24: 1513–1521.
29. Mallat S. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. Pattern Recognition Machine Intell.* 1989; 11: 674–693.
30. Conti G, Tambalo S, Villetti G, Catinella S, Carnini C, Bassani F, Sonato N, Sbarbati A, Marzola P. Evaluation of lung inflammation induced by intratracheal administration of LPS in mice: comparison between MRI and histology. *MAGMA* 2010; 23: 93–101.
31. Beckmann N, Tigani B, Ekatodramis D, Borer R, Mazzoni L, Fozard JR. Pulmonary edema induced by allergen challenge in the rat: noninvasive assessment by magnetic resonance imaging. *Magn. Reson. Med.* 2001; 45: 88–95.
32. Ble FX, Cannet C, Zurbrugg S, Karmouty-Quintana H, Bergmann R, Frossard N, Trifilieff A, Beckmann N. Allergen-induced lung inflammation in actively sensitized mice assessed with MR imaging. *Radiology* 2008; 248: 834–843.
33. Dehnert C, Risse F, Ley S, Kuder TA, Buhmann R, Puderbach M, Menold E, Mereles D, Kauczor HU, Bartsch P, Fink C. Magnetic resonance imaging of uneven pulmonary perfusion in hypoxia in humans. *Am. J. Respir. Crit. Care Med.* 2006; 174: 1132–1138.
34. De Becker J, Bister M, Langlois N, Vanhove C, Demonceau G, Cornelis J. A split-and-merge algorithm for the segmentation of 2-D, 3-D, 4-D cardiac images. *Proceedings of the IEEE Satellite Symposium on 3D Advanced Image Processing in Medicine*, Rennes, France, 1992; 185–189.