

Pulmonary Ventilation Maps Generated with Free-breathing Proton MRI and a Deep Convolutional Neural Network

Dante P. I. Capaldi, PhD* • Fumin Guo, PhD* • Lei Xing, PhD • Grace Parraga, PhD, FCAHS

From the Department of Radiation Oncology, School of Medicine, Stanford University, Stanford, Calif (D.P.I.C., L.X.); Sunnybrook Research Institute, Department of Medical Biophysics, University of Toronto, Toronto, Canada (F.G.); and Robarts Research Institute, Department of Medical Biophysics, The University of Western Ontario, 1151 Richmond St N, London, ON, Canada N6A 5B7 (G.P.). Received June 25, 2020; revision requested July 29; revision received October 17; accepted October 27. **Address correspondence** to G.P. (e-mail: gparraga@robarts.ca).

D.P.I.C. and F.G. are supported by postdoctoral fellowships from the Natural Sciences and Engineering Research Council of Canada. L.X. is supported by the National Institutes of Health (grant nos. 1R01CA223667 and 1R01CA227713). G.P. is supported by the Canada Research Chairs Program and holds a Tier 1 Canada Research Chair position.

* D.P.I.C. and F.G. contributed equally to this work.

Conflicts of interest are listed at the end of this article.

See also the editorial by Vogel-Claussen in this issue.

Radiology 2021; 298:427–438 • <https://doi.org/10.1148/radiol.2020202861> • Content codes: **CH MR**

Background: Hyperpolarized noble gas MRI helps measure lung ventilation, but clinical translation remains limited. Free-breathing proton MRI may help quantify lung function using existing MRI systems without contrast material and may assist in providing information about ventilation not visible to the eye or easily extracted with segmentation methods.

Purpose: To explore the use of deep convolutional neural networks (DCNNs) to generate synthetic MRI ventilation scans from free-breathing MRI (deep learning [DL] ventilation MRI)—derived specific ventilation maps as a surrogate of noble gas MRI and to validate this approach across a wide range of lung diseases.

Materials and Methods: In this secondary analysis of prospective trials, 114 paired noble gas MRI and two-dimensional free-breathing MRI scans were obtained in healthy volunteers with no history of chronic or acute respiratory disease and in study participants with a range of different obstructive lung diseases, including asthma, bronchiectasis, chronic obstructive pulmonary disease, and non-small-cell lung cancer between September 2013 and April 2018 (ClinicalTrials.gov identifiers: NCT03169673, NCT02351141, NCT02263794, NCT02282202, NCT02279329, and NCT02002052). A U-Net–based DCNN model was trained to map free-breathing proton MRI to hyperpolarized helium 3 (^3He) MRI ventilation and validated using a sixfold validation. During training, the DCNN ventilation maps were compared with noble gas MRI scans using the Pearson correlation coefficient (r) and mean absolute error. DCNN ventilation images were segmented for ventilation and ventilation defects and were compared with noble gas MRI scans using the Dice similarity coefficient (DSC). Relationships were evaluated with the Spearman correlation coefficient (r_s).

Results: One hundred fourteen study participants (mean age, 56 years \pm 15 [standard deviation]; 66 women) were evaluated. As compared with ^3He MRI, DCNN model ventilation maps had a mean r value of 0.87 ± 0.08 . The mean DSC for DL ventilation MRI and ^3He MRI ventilation was 0.91 ± 0.07 . The ventilation defect percentage for DL ventilation MRI was highly correlated with ^3He MRI ventilation defect percentage ($r_s = 0.83$, $P < .001$, mean bias = $-2.0\% \pm 5$). Both DL ventilation MRI ($r_s = -0.51$, $P < .001$) and ^3He MRI ($r_s = -0.61$, $P < .001$) ventilation defect percentage were correlated with the forced expiratory volume in 1 second. The DCNN model required approximately 2 hours for training and approximately 1 second to generate a ventilation map.

Conclusion: In participants with diverse pulmonary pathologic findings, deep convolutional neural networks generated ventilation maps from free-breathing proton MRI trained with a hyperpolarized noble-gas MRI ventilation map data set. The maps showed correlation with noble gas MRI ventilation and pulmonary function measurements.

© RSNA, 2020

Ventilation heterogeneity is a hallmark finding in obstructive lung disease, including asthma and chronic obstructive pulmonary disease (COPD), and it reflects the underlying pathophysiologic airway and lung parenchymal abnormalities (1,2). Unfortunately, the structure-function relationship between airway obstruction and parenchymal destruction with ventilation heterogeneity, albeit well recognized, is complex and not well understood. This knowledge gap contributes to poor disease control and a greater incidence of acute disease worsening or exacerbations (3). As the cost and frequency of hospitalizations continue to rise

(4), investigations into quantitative markers of ventilation heterogeneity may further our understanding of the complex relationships between lung structure and function with disease control in obstructive lung disease.

Pulmonary imaging methods, such as CT and MRI, can regionally help visualize lung disease (5,6). High-spatial-resolution CT is the clinical imaging tool of choice for visualizing and quantifying lung structural abnormalities, such as airway and emphysematous tissue destruction (7,8). Complementary functional information, specifically regional ventilation heterogeneity, uses inhaled contrast-enhanced

Abbreviations

COPD = chronic obstructive pulmonary disease, DCNN = deep convolutional neural network, DL = deep learning, DSC = Dice similarity coefficient, FEV₁ = forced expiratory volume in 1 second, MAE = mean absolute error, NSCLC = non–small-cell lung cancer, 2D = two-dimensional, VDP = ventilation defect percentage

Summary

In participants with diverse pulmonary pathologic findings, a deep learning approach generated pulmonary ventilation maps from free-breathing proton MRI that showed a strong correlation with noble gas MRI ventilation and pulmonary function measurements.

Key Results

- A deep learning (DL) approach to generating pulmonary ventilation maps from free-breathing proton MRI showed a strong correlation with hyperpolarized noble gas MRI ventilation defect percentage ($r_s = 0.83$, $P < .001$).
- Both DL ($r_s = -0.51$, $P < .001$) and noble gas MRI ($r_s = -0.61$, $P < .001$) ventilation defect percentage measurements correlated with the forced expiratory volume in 1 second, and these correlation coefficients were not different ($P = .28$).
- The DL model, once trained, generated a predicted lung ventilation map within 1 second.

imaging techniques, such as hyperpolarized noble gas MRI (9), dual-energy xenon CT (10), and nuclear imaging (11,12). Alternative approaches, such as four-dimensional CT (13), Fourier decomposition MRI (14,15), and four-dimensional MRI (16), can generate ventilation maps without exogenous contrast material. These approaches leverage dynamic structural imaging and quantify lung ventilation by measuring differences in signal intensities or local tissue deformation throughout the imaging series. Despite the unique advantages, validation remains a concern because of the numerous image acquisition protocols and algorithms developed to measure regional ventilation. Disagreements between methods (17) and the finding of only modest relationships when comparing unenhanced and contrast-based ventilation imaging methods (18) have prompted several comprehensive multi-institutional validation studies (19).

With the recent advances in deep learning (DL) using convolutional neural networks, groundbreaking performance has been demonstrated in various medical imaging applications, including image segmentation and classification (20). A recent study (21) employed a deep convolutional neural network (DCNN) to model the complex relationship between CT and MRI scans, namely image translation or synthesis, to generate synthetic CT scans as a way to simplify the workflow for MRI-guided radiation therapy. Accordingly, our purpose was to explore the use of DCNN to generate synthetic MRI ventilation scans from free-breathing proton MRI–derived specific ventilation maps as a surrogate of hyperpolarized noble gas MRI and to validate this approach across a wide range of lung diseases.

Materials and Methods

Study Participants and Design

In this secondary analysis of a prospective trial, we evaluated participants with different lung diseases, including

asthma, bronchiectasis, COPD, and non–small cell lung cancer (NSCLC), as well as healthy volunteers. All participants provided written informed consent to a research ethics board (institutional ethics board no. 00000984). The study was approved and compliant with the Health Insurance Portability and Accountability Act in registered studies (ClinicalTrials.gov identifiers: NCT03169673 for healthy volunteers, NCT02351141 and NCT02263794 for participants with asthma, NCT02282202 for participants with bronchiectasis; NCT02279329 and NCT02282202 for participants with COPD; and NCT02002052 for participants with NSCLC). Participants aged 18–85 years were recruited from local tertiary care clinics and through advertisements between September 2013 and April 2018 on the basis of the inclusion and exclusion criteria. For all participants, the exclusion criteria were as follows: (a) serious claustrophobia, (b) unable to undergo spirometry, (c) body mass index greater than 40 kg/m², and (d) any contraindications to MRI (ie, metal, electronic, or magnetic implants). Inclusion criteria for healthy volunteers were as follows: (a) medical history with no chronic or acute respiratory disease and stable health, (b) age of 18–75 years, and (c) smoking history of less than 1 pack-year. Inclusion criteria for participants with asthma were as follows: (a) clinical diagnosis of asthma with β_2 -agonist reversibility of forced expiratory volume in 1 second (FEV₁) greater than 12% or methacholine FEV₁ provocative concentration that decreased FEV₁ by 20% (PC₂₀) (≤ 8 mg/mL if not administered inhaled corticosteroid or ≤ 16 mg/mL if administered inhaled corticosteroid), (b) no history of other chronic disease, (c) age of 18–85 years, and (d) smoking history of less than 1 pack-year. Inclusion criteria for participants with bronchiectasis were as follows: (a) age of 40–85 years and (b) clinical diagnosis of bronchiectasis and chronic mucous production identified. Inclusion criteria for participants with COPD were as follows: (a) age of 40–85 years and (b) clinical diagnosis of COPD or a smoking history of greater than 10 pack-years. Inclusion criteria for participants with NSCLC were as follows: (a) age older than 18 years, (b) histologically confirmed unresectable stage IIIA or IIIB NSCLC, (c) a smoking history of greater than 10 pack-years, (d) an Eastern Cooperative Oncology Group performance score between 0 and 2, (e) FEV₁ greater than 30% percentage predicted, and (f) able to undergo platinum-based chemotherapy as determined by his or her treating physicians. For participants with NSCLC, exclusion criteria also included (a) history of lung cancer within 5 years, (b) serious medical comorbidities (eg, unstable angina or sepsis) or other contraindications to radiation therapy or chemotherapy, (c) previous thoracic radiation at any time, (d) metastatic disease, (e) inability to attend the full course of radiation therapy or follow-up visits, and (f) pregnant or lactating women.

Seventy-two of the 114 study participants had been previously reported (15,16,22). The previous work addressed determining the feasibility of using Fourier decomposition and specific ventilation free-breathing proton MRI in the evaluation of ventilation defects in different participant groups. Alternatively, this study focused on the development of using DL to predict ventilation and ventilation defects.

Pulmonary Function Tests

Spirometry and plethysmography were performed according to American Thoracic Society and European Respiratory Society guidelines (23,24) using a body plethysmograph (Elite Series Plethysmograph; MGC Diagnostics) to obtain the FEV₁, forced vital capacity, residual volume, and total lung capacity.

MRI Examination

All MRI examinations were performed in the coronal plane with a whole-body 3.0-T system (Discovery MR750; GE Healthcare). Each participant underwent free-breathing hydrogen 1 (¹H) MRI, static breath-hold ¹H MRI, and hyperpolarized helium 3 (³He) MRI (15).

Static breath-hold ¹H MRI was performed with a whole-body radiofrequency coil and a fast spoiled gradient-recalled-echo sequence with the following parameters: total acquisition time, 16 seconds; repetition time msec/echo time msec, 4.7/1.2; flip angle, 30°; field of view, 40 × 40 cm²; matrix, 128 × 80 (zero-padded to 128 × 128); bandwidth, 24.4 kHz; 15–17 sections; and section thickness, 15 mm without gap. Static breath-hold hyperpolarized ³He MRI was performed with a single-channel, rigid elliptical transmit-receive chest coil (Rapid Biomedical) and a fast spoiled gradient-recalled-echo sequence with the following parameters: total acquisition time, 14 seconds; 4.3/1.4; flip angle, 7°; field of view, 40 × 40 cm²; bandwidth, 48.8 kHz; matrix, 128 × 80 (zero-padded to 128 × 128); 15–17 sections; section thickness, 15 mm without gap. A commercial polarizer (HeliSpin; Polarean Imaging) was used to polarize ³He to 30%–40%, where participants inhaled a 1-L gas mixture (30% hyperpolarized ³He and 70% medical-grade N₂ gas by volume) after passive expiration to functional residual capacity. Free-breathing ¹H MRI was performed using a 32-channel torso coil (GE Healthcare) and an optimized balanced steady-state free-precession sequence with the following parameters: total acquisition time, 125 seconds; 1.9/0.6; flip angle, 15°; field of view, 40 × 40 cm²; bandwidth, 250 kHz; matrix, 256 × 256; number of phases, 500; one section; section thickness, 15 mm without gap; and temporal resolution, 0.25 second. Respiratory bellows were used to monitor respiratory rate.

Image Analysis

Free-breathing ¹H MRI and ³He MRI scans were analyzed as previously described (16,25), with Matlab software (version R2019a; Mathworks). Briefly, the original free-breathing series was sorted to a single respiratory phase according to the magnitude of the corresponding respiratory bellow data. The sorted image series was interpolated and registered to a single phase using a modality-independent neighborhood descriptor-based deformable registration (26), creating a data set with respiratory and cardiac motion compensated, where the reference phase was tidal inspiration. The coregistered tidal inspiration and expiration volumes were used to generate specific ventilation maps as shown in Figure 1. Briefly, specific ventilation (SV) was determined on a voxel by voxel basis using the following equation (16): $SV \approx (SI_{exp} - SI_{insp})/SI_{insp}$, where SI_{exp} and SI_{insp} are the signal intensities at expiration and inspiration, respectively. The reference tidal inspiration

image was segmented using a multiregional segmentation approach, as previously described (27). The specific ventilation map within the segmented lung mask was extracted before DCNN implementation.

Training, Validation, and Testing

Before DCNN training and testing, ³He MRI scans and the masked specific ventilation map were coregistered (16). We selected corresponding two-dimensional (2D) static ¹H and/or ³He MRI scans that were at the same anteroposterior coronal slice to coregister to the 2D free-breathing MRI ventilation map (15). The selected 2D ³He MRI scan was coregistered to the reference tidal inspiration image, facilitated by first registering the 2D ³He MRI scan to the selected corresponding 2D ¹H MRI scan, by means of landmark registration (25), then coregistering the 2D ¹H MRI scan to the reference tidal inspiration image using a modality-independent neighborhood descriptor (26). Finally, the specific ventilation and ³He MRI scans were scaled to 0–255 signal intensities and a 256 × 256 matrix size. Histogram matching was performed to help standardize image intensities across different participants. The generated specific ventilation maps and the registered ³He MRI scans were used to train the DCNN. The DCNN model was generated based on a U-Net architecture and comprised convolutions with rectified linear activation, max pooling, up pooling, and skip connections between the two branches. The depth of the feature map and image size from each layer are provided in Figure 1. During training, the model prediction DL ventilation MRI was backpropagated to compare against the ground truth (registered ³He MRI) to update DCNN parameters.

DCNN training and testing were performed with TensorFlow 1.14 (<https://www.tensorflow.org/>) and Python 3.6 (<https://www.python.org/>) platforms on graphics processing unit (Titan RTX; NVIDIA). The code is available online from the authors (<https://github.com/capalddid/DLMRI>). Figure 1 illustrates the overall architecture of the 2D DCNN model (21). The DCNN was built based on the U-Net architecture (28) and comprised two branches—an encoding and a decoding path. Each layer in the two paths comprises a 3 × 3 convolution, a rectified linear unit to introduce nonlinearity, max pooling and/or up pooling, and skip connections between the two branches at the same resolution level. The encoder is similar to a traditional convolutional neural network and used to extract a hierarchy of increasingly complex features from an input-specific ventilation map. The decoding part transforms the features extracted from the encoding branch and gradually reconstructs a synthetic image. The final output of the network is a 2D model prediction, coined DL ventilation MRI, with the same size as the input image. To optimize the network parameters, the mean absolute error (MAE) between the predicted ventilation image and corresponding ³He MRI scan was used as the loss function. MAE was used as the loss function as it is more robust to outliers in the training data, such as noise or other artifacts in the images or because of imperfect matching between images (21). Backpropagation (29) with an Adam solver at a learning rate of 10^{−4} was implemented to train the DCNN model.

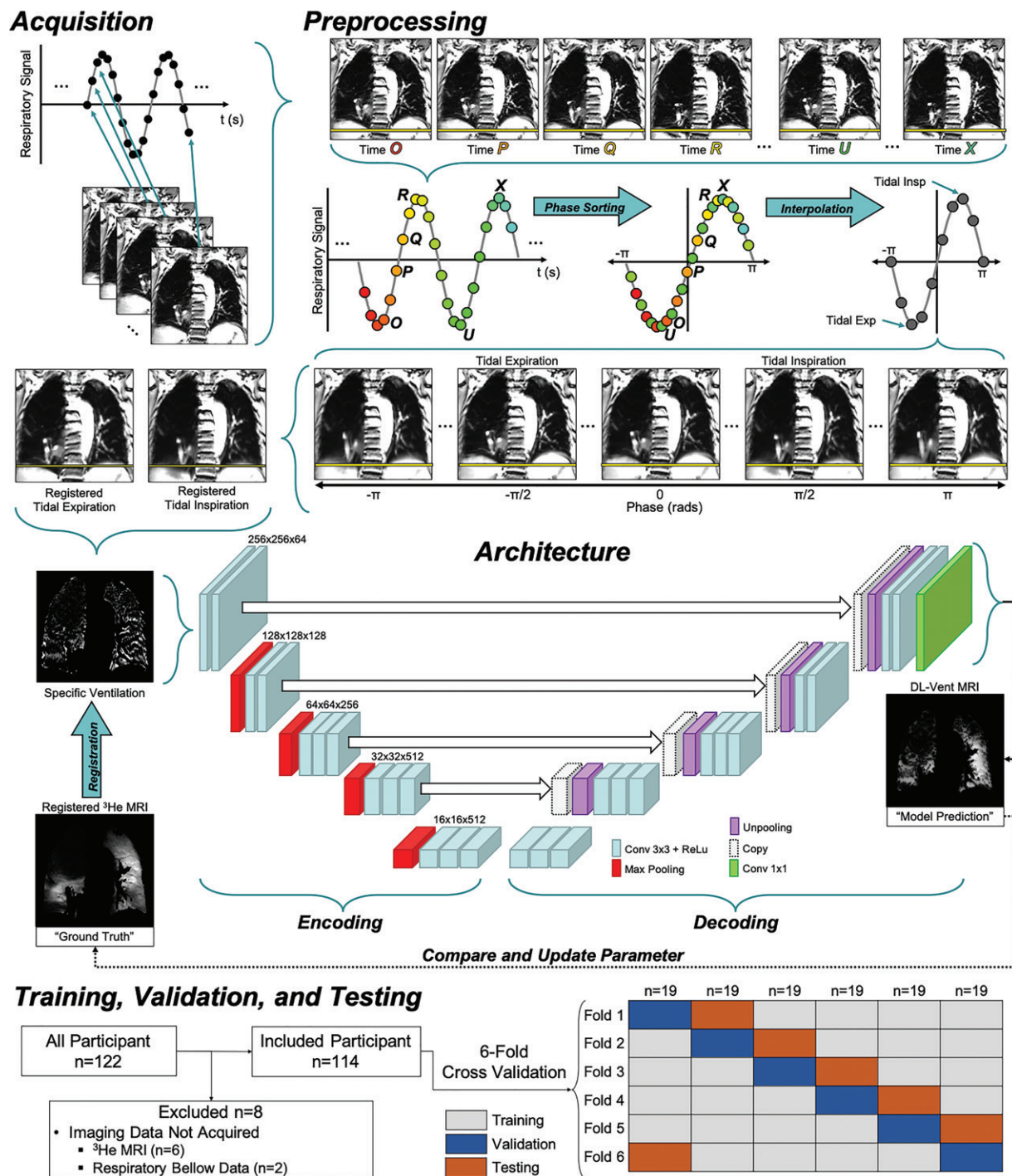


Figure 1: Free-breathing hydrogen 1 (^1H) MRI pipeline for generating synthetic ventilation images. Acquisition: Free-breathing ^1H MRI scans were obtained with corresponding respiratory bellow data trace during multiple (variable) breathing cycles. Preprocessing: colors are associated with time that each image (time O–X) is obtained and correspond to color points on respiratory bellow trace. Respiratory phases over time were sorted into single section ($-\pi$ to π). Original image series was sorted according to order of respiratory phases and interpolated to generate respiratory phase-binned images and tidal expiration and tidal inspiration frames. Yellow line on tidal inspiration images indicates level of diaphragm. Tidal expiration images were registered to tidal inspiratory frames to generate specific ventilation maps. Generated specific ventilation maps and registered helium 3 (^3He) MRI scans were used to train deep convolutional neural network (DCNN). Architecture: The DCNN model was generated based on U-Net architecture and comprises convolutions with rectified linear activation (blue), max pooling (red), up pooling (purple), and skip connections (arrow) between the two branches. Depth of feature map and image size from each layer are provided. During training, model prediction (deep learning ventilation MRI (DL-Vent MRI) was backpropagated to compare against the ground truth (registered ^3He MRI) to update DCNN parameters. Training, Validation, and Testing: Flowchart shows participants enrolled in study. For training, validation, and testing, participants were randomly divided into six groups. Training was performed during six iterations, whereby for each iteration, the model was trained on four groups (gray), validated on one group (blue), and tested on one group (orange). Conv = convolution Exp = expiration, Insp = inspiration, rads = radians, t(s) = time in seconds.

Table 1: Participant Demographic Information, Results of Pulmonary Function Tests, and Imaging Measurements

Parameter	All (n = 114)	Healthy (n = 7)	Asthma (n = 60)	Bronchiectasis (n = 14)	COPD (n = 16)	NSCLC (n = 17)	P Value
Demographic information							
Age (y)	56 ± 15	26 ± 3	50 ± 10	70 ± 11	65 ± 9	68 ± 10	<.001*
No. of women	66	3	35	10	8	10	...
BMI (kg/m ²)	26 ± 5	23 ± 4	28 ± 4	23 ± 4	26 ± 4	26 ± 4	.007 [†]
Pulmonary function test results							
FEV ₁ (% _{pred})	72 ± 22	99 ± 7	72 ± 20	68 ± 22	60 ± 23	77 ± 23	.002*
FEV ₁ /FVC (%)	66 ± 15 [‡]	82 ± 5	66 ± 14 [§]	70 ± 12	51 ± 14	68 ± 12	<.001 [†]
RV (% _{pred})	130 ± 32 [‡]	104 ± 18	129 ± 27 [§]	128 ± 30	155 ± 38	122 ± 37	.02*
RV/TLC (%)	44 ± 11 [‡]	25 ± 4	41 ± 8 [§]	54 ± 11	50 ± 10	47 ± 9	<.001*
Imaging measurements							
³ He MRI VDP (%)	14 ± 10	5 ± 3	11 ± 7	23 ± 11	24 ± 11	14 ± 8	<.001 [†]
DL ventilation MRI VDP (%)	12 ± 9	5 ± 1	9 ± 6	20 ± 12	22 ± 11	13 ± 7	<.001 [†]

Note.—Except where indicated, data are means ± standard deviations. BMI = body mass index, COPD = chronic obstructive pulmonary disease, DL = deep learning, FEV₁ = forced expiratory volume in 1 second, FVC = forced vital capacity, ³He = helium 3, NSCLC = non-small-cell lung cancer, %_{pred} = percentage of predicted value, RV = residual volume, TLC = total lung capacity, VDP = ventilation defect percentage.

* The significant difference between subgroups ($P < .05$) was determined with analysis of variance for parametric data.

[†] The significant difference between subgroups ($P < .05$) was determined with Kruskal-Wallis tests for nonparametric data and adjusted with the Holm-Bonferroni correction.

[‡] $n = 111$ participants.

[§] $n = 57$ participants.

A sixfold cross-validation was used to evaluate the performance of the DCNN model. The image pairs (114 in total) were randomly and evenly divided into six groups (Group₁, Group₂, Group₃, Group₄, Group₅, and Group₆; each with 19 image pairs). For each fold, one group (Group_{*i*}) was held back for testing, and the remaining groups (Group_{*i+1*}) were used for training the DCNN ($i = 1, 2, 3, 4, 5, 6$). Thus, no information is carried over from training to testing. For each session, the DCNN was trained for 300 epochs, where the MAE was calculated with the training data, and the model with the lowest MAE during training was saved and applied to the test data set. The trained model was applied to the test free-breathing ¹H MRI-specific ventilation maps to generate DL ventilation MRI scans.

Algorithm Performance Evaluation

To evaluate the performance of the proposed DCNN model, DL ventilation MRI maps (MRI_{DL}) were compared with the hyperpolarized noble gas MRI scans (MRI_{NG}) on a voxel-by-voxel basis using MAE, as follows:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |\text{MRI}_{\text{NG}}(i) - \text{MRI}_{\text{DL}}(i)|,$$

where n is the total number of voxels inside the lung region. The relative mean absolute error, relative mean error, and Pearson correlation coefficients were also used to compare DL and noble gas MRI ventilation signals on a voxel-wise basis. The DL ventilation MRI and ³He ventilation MRI scans were segmented for ventilation and defect to quantify the ventilation

defect percentage (VDP) using custom software generated with Matlab (version R2019a; Mathworks), as previously described (16,25). Briefly, the ventilation defect volume was determined after performing hierarchical κ -means clustering of the ventilation image (either ³He or DL) and segmenting the ¹H MRI scans automatically using a seeded region-growing algorithm. The Dice similarity coefficient (DSC) was used to quantify the spatial overlap between DL ventilation MRI and ³He MRI ventilation and ventilation defect volumes (30). DSC evaluates the overlap ratio between two regions, R_a and R_b , and is computed as follows:

$$\text{DSC} = \frac{2|R_a \cap R_b|}{|R_a| + |R_b|} \times 100\%,$$

where \cap represents the intersection between two binary (ventilation and/or defects) regions and $|\cdot|$ denotes the size of a region R .

Statistical Analysis

Shapiro-Wilk tests were used to determine the normality of the data distribution. Nonparametric tests were performed for data that were determined to be not normally distributed. Between-group differences were determined using a one-way analysis of variance and a Kruskal-Wallis test for parametric and nonparametric data, respectively, and adjusted with the Holm-Bonferroni correction. Differences between DL ventilation MRI and ³He MRI ventilation measurements were determined with Wilcoxon matched pairs signed-rank tests. Relationships between ventilation imaging measurements and pulmonary function tests were

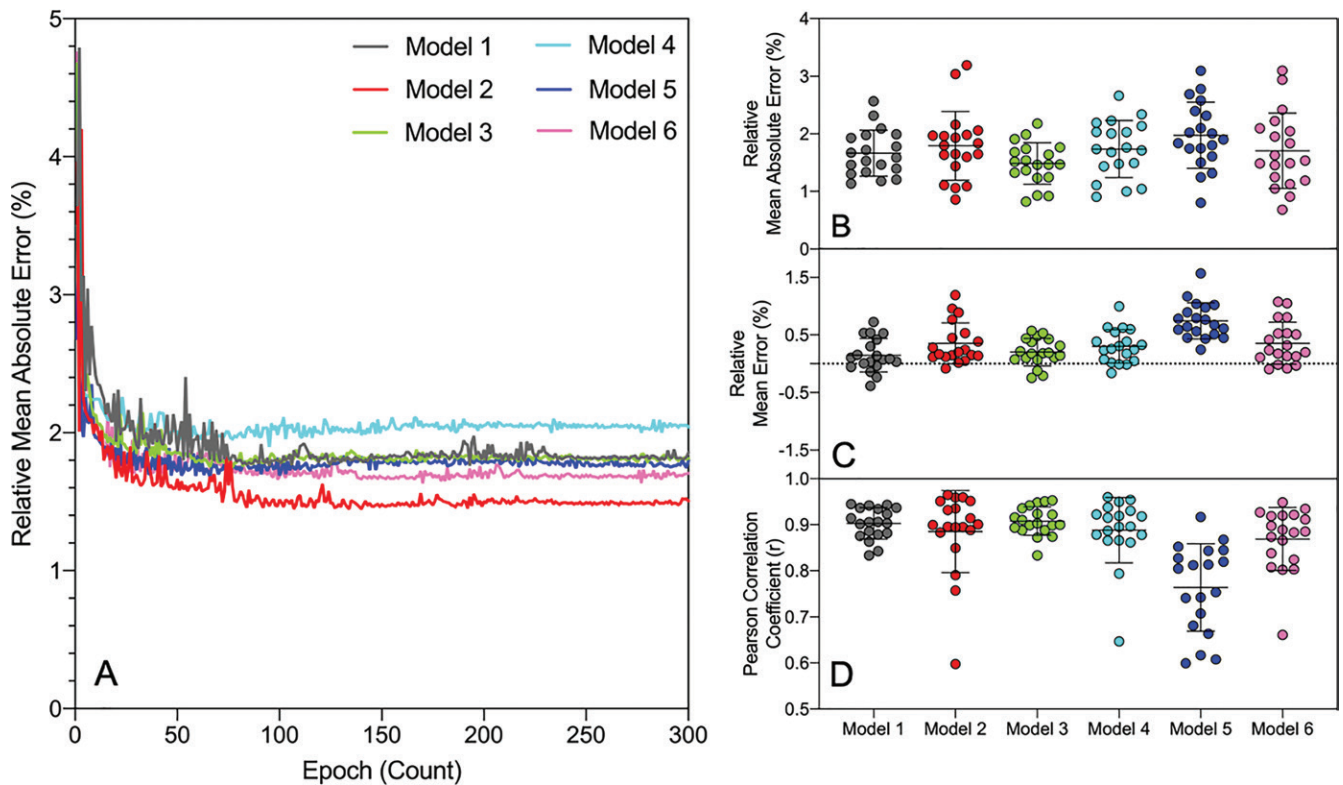


Figure 2: Cross-validation training and evaluation of the deep convolutional neural network (DCNN) model. Entire data set ($n = 114$) was evenly divided into six groups (Group₁, Group₂, Group₃, Group₄, Group₅, and Group₆; each with $n = 19$) for sixfold cross-validation. For each fold, Group_{*i*} and Group_{*i*+1} were used for DCNN training and validation, respectively ($i = 1, 2, 3, 4, 5, 6$), and the remaining groups were used for testing. A, Graph shows loss function, using mean absolute error, as function of number of epochs during training. B–D, Sixfold cross-validation on the test data set evaluated using means \pm standard deviations of, B, relative mean absolute error (model 1 = $1.7\% \pm 0.4$; model 2 = $1.8\% \pm 0.6$; model 3 = $1.5\% \pm 0.4$; model 4 = $1.7\% \pm 0.5$; model 5 = $2.0\% \pm 0.6$; model 6 = $1.7\% \pm 0.7$), C, relative mean error (model 1 = $-0.1\% \pm 0.3$; model 2 = $0.4\% \pm 0.4$; model 3 = $0.2\% \pm 0.2$; model 4 = $0.3\% \pm 0.3$; model 5 = $0.7\% \pm 0.3$; model 6 = $0.4\% \pm 0.4$), and, D, Pearson correlation coefficient (model 1 = 0.90 ± 0.03 ; model 2 = 0.89 ± 0.09 ; model 3 = 0.91 ± 0.03 ; model 4 = 0.88 ± 0.07 ; model 5 = 0.76 ± 0.10 ; model 6 = 0.87 ± 0.07) for each model.

determined with Spearman coefficients (r_s). Correlation coefficients were compared with the Fisher z transformation (31). The Bland-Altman method (32) was used to evaluate agreement between DL ventilation MRI and ^3He MRI ventilation maps. Results were considered significant when the probability of two-tailed type I error (α) was less than 5% ($P < .05$). Statistical analysis was performed using software (GraphPad Prism, version 8.1.2; GraphPad Software).

Results

Participant Demographics

Of the 122 participants enrolled in the study, eight were excluded for not having ^3He MRI data ($n = 6$) or respiratory bellow data used for preprocessing the free-breathing MRI ($n = 2$) (Fig 1). Table 1 shows participant demographic information, pulmonary function test results, and imaging measurements for all participants. One hundred fourteen participants (mean age, $56 \text{ years} \pm 15$; 66 women) were evaluated, including seven healthy volunteers with no history of chronic or acute respiratory disease (mean age, $26 \text{ years} \pm 3$; four women), 60 participants with asthma (mean age, $50 \text{ years} \pm 10$; 35 men), 14 participants with bronchiectasis (mean age, $70 \text{ years} \pm 11$; 10 women), 16 partici-

pants with COPD (mean age, $65 \text{ years} \pm 9$; eight women), and 17 participants with NSCLC (mean age, $68 \text{ years} \pm 10$; 10 women).

Loss and Accuracy of the DCNN Model during Training

Figure 2 shows the loss and the accuracy of the DCNN model during training. Across all six folds, the DCNN model produced ventilation maps with a mean relative MAE of $1.7\% \pm 0.5$, relative mean error of $0.4\% \pm 0.4$, and an average Pearson correlation coefficient of 0.87 ± 0.08 when compared with the ^3He MRI. Initial DCNN training required approximately 20 minutes per model (approximately 2 hours in total for the six-fold cross-validation), and applying the trained DCNN model to a test data set to generate a synthetic ventilation map required approximately 1 second.

Qualitative and Quantitative Agreement between DL Ventilation MRI and ^3He MRI

Figure 3 shows qualitative spatial agreement between the DL ventilation MRI and ^3He MRI ventilation scans for five participants from all subgroups. Figure 4 shows the quantitative agreement between DL ventilation MRI and ^3He MRI ventilation measurements. Across all participant subgroups except for those with asthma ($P < .001$), DL ventilation MRI and

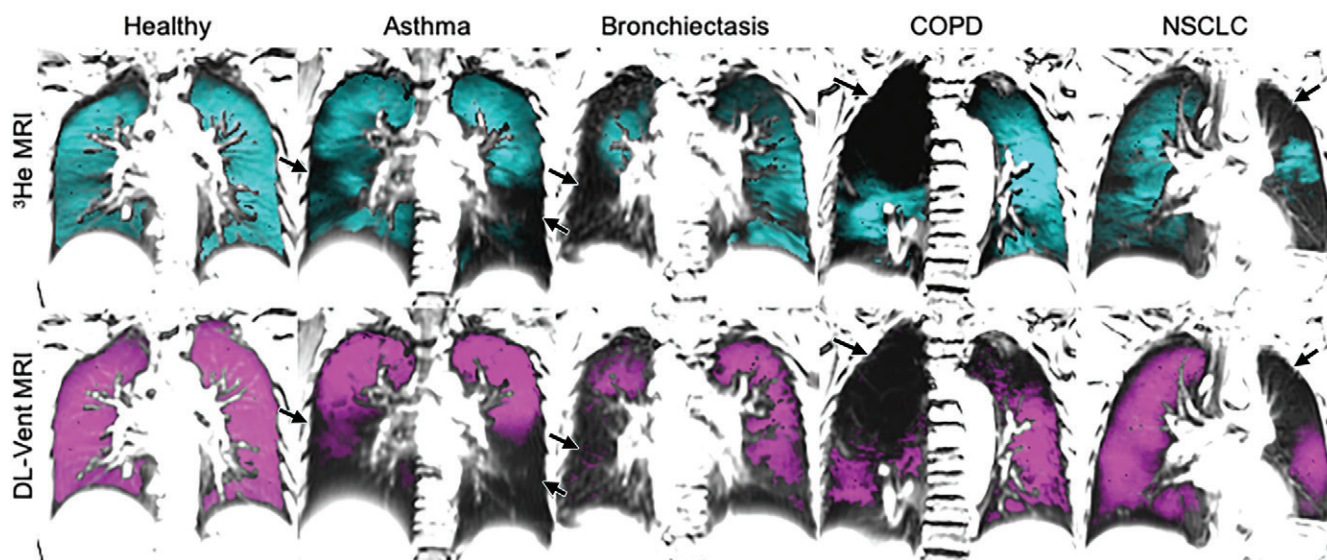


Figure 3: Qualitative comparison of deep learning (DL) ventilation MRI (DL-Vent MRI) versus helium 3 (^3He) MRI ventilation. Hyperpolarized ^3He MRI (cyan) and DL ventilation MRI (magenta) ventilation images are coregistered to anatomic ^1H MRI (gray scale) for multiple participants, including a healthy volunteer (31-year-old woman; forced expiratory volume in 1 second [FEV_1], 102% percentage of predicted value; ^3He MRI ventilation defect percentage [VDP], 5%; DL ventilation MRI VDP, 5%), a participant with asthma (41-year-old woman; FEV_1 , 60% percentage of predicted value; ^3He MRI VDP, 24%; DL ventilation MRI VDP, 35%), a participant with bronchiectasis (58-year-old woman; FEV_1 , 43% percentage of predicted value; ^3He MRI VDP, 29%; DL ventilation MRI VDP, 37%), a participant with chronic obstructive pulmonary disease (COPD) (80-year-old man; FEV_1 , 77% percentage of predicted value; ^3He MRI VDP, 40%; DL ventilation MRI VDP, 45%), and a participant with non-small-cell lung cancer (NSCLC) (44-year-old man; FEV_1 , 61% percentage of predicted value; ^3He MRI VDP, 26%; DL ventilation MRI VDP, 22%). Arrows indicate corresponding regions of ventilation defect visible across both imaging methods.

^3He MRI VDP were not different (healthy participants, $P > .99$; participants with bronchiectasis, $P > .99$; participants with COPD, $P = .59$; participants with NSCLC, $P > .99$). For the entire data set, DL ventilation MRI VDP was highly correlated with ^3He MRI VDP ($r_s = 0.83$, $P < .001$), with a slope close to unity (slope = 0.87). Bland-Altman analysis of agreement for DL ventilation MRI and ^3He MRI VDP demonstrated a marginal bias of $-2.0\% \pm 5$.

Spatial Correspondence of DL Ventilation MRI and ^3He MRI Ventilation and/or Defects

Figure 5 shows the spatial correspondence of DL ventilation MRI and ^3He MRI ventilation and/or defects for the participants shown in Figure 3. Table 2 shows that the mean DSC for DL ventilation MRI and ^3He MRI ventilation was $90\% \pm 6$. For ventilation defects, the mean DSC was $37\% \pm 14$ across all participants. Upon investigating these spatial relationships further, we observed relationships between DSC for DL ventilation MRI and ^3He MRI ventilation with VDP (^3He MRI VDP, $r_s = -0.95$, $P < .001$; DL ventilation MRI VDP, $r_s = -0.86$, $P < .001$) and between DSC for DL ventilation MRI and ^3He MRI defects with VDP (^3He MRI VDP, $r_s = 0.47$, $P < .001$; DL ventilation MRI VDP, $r_s = 0.63$, $P < .001$).

Pulmonary Function Tests Showing DL Ventilation MRI and ^3He MRI VDP Correlation

Figure 6 and Table 3 summarize relationships for DL ventilation MRI and ^3He MRI VDP with pulmonary function measurements. There were significant correlations for DL ventilation MRI and ^3He MRI VDP with FEV_1 (DL ventilation MRI, $r_s = -0.51$, $P < .001$; ^3He MRI, $r_s = -0.61$, $P < .001$),

$\text{FEV}_1/\text{forced vital capacity}$ (DL ventilation MRI, $r_s = -0.45$, $P < .001$; ^3He MRI, $r_s = -0.57$, $P < .001$), residual volume (DL ventilation MRI, $r_s = 0.36$, $P < .001$; ^3He MRI, $r_s = 0.50$, $P < .001$), and residual volume/total lung capacity (DL ventilation MRI, $r_s = 0.60$, $P < .001$; ^3He MRI, $r_s = 0.60$, $P < .001$). Pearson correlations for DL ventilation MRI and ^3He MRI VDP were not different across all pulmonary function tests, as shown in Table 3.

Discussion

Hyperpolarized noble gas MRI measures lung ventilation, albeit clinical translation remains limited. Alternatively, free-breathing proton MRI may help quantify lung function using existing MRI systems without contrast material and may provide information about ventilation not visible to the eye nor easily extracted with segmentation methods. We generated a deep convolutional neural network model to create synthetic ventilation images from free-breathing proton MRI. For a divergent but relatively small group of 114 participants, we made the following observations: (a) deep learning (DL) ventilation MRI ventilation defect percentage (VDP) measurements were correlated with helium 3 (^3He) MRI VDP measurements ($r_s = 0.83$, $P < .001$); (b) DL ventilation MRI scans were correlated with ^3He MRI scans (Dice similarity coefficient = $90\% \pm 6$); and (c) both DL ventilation MRI VDP and ^3He MRI VDP correlated with forced expiratory volume in 1 second with similar correlation coefficients (DL ventilation MRI, $r_s = -0.51$, $P < .001$; ^3He MRI, $r_s = -0.61$, $P < .001$).

The structural abnormalities that manifest as ventilation heterogeneity have been determined with multimodality pulmonary imaging (5,6), which has prompted the development of complex

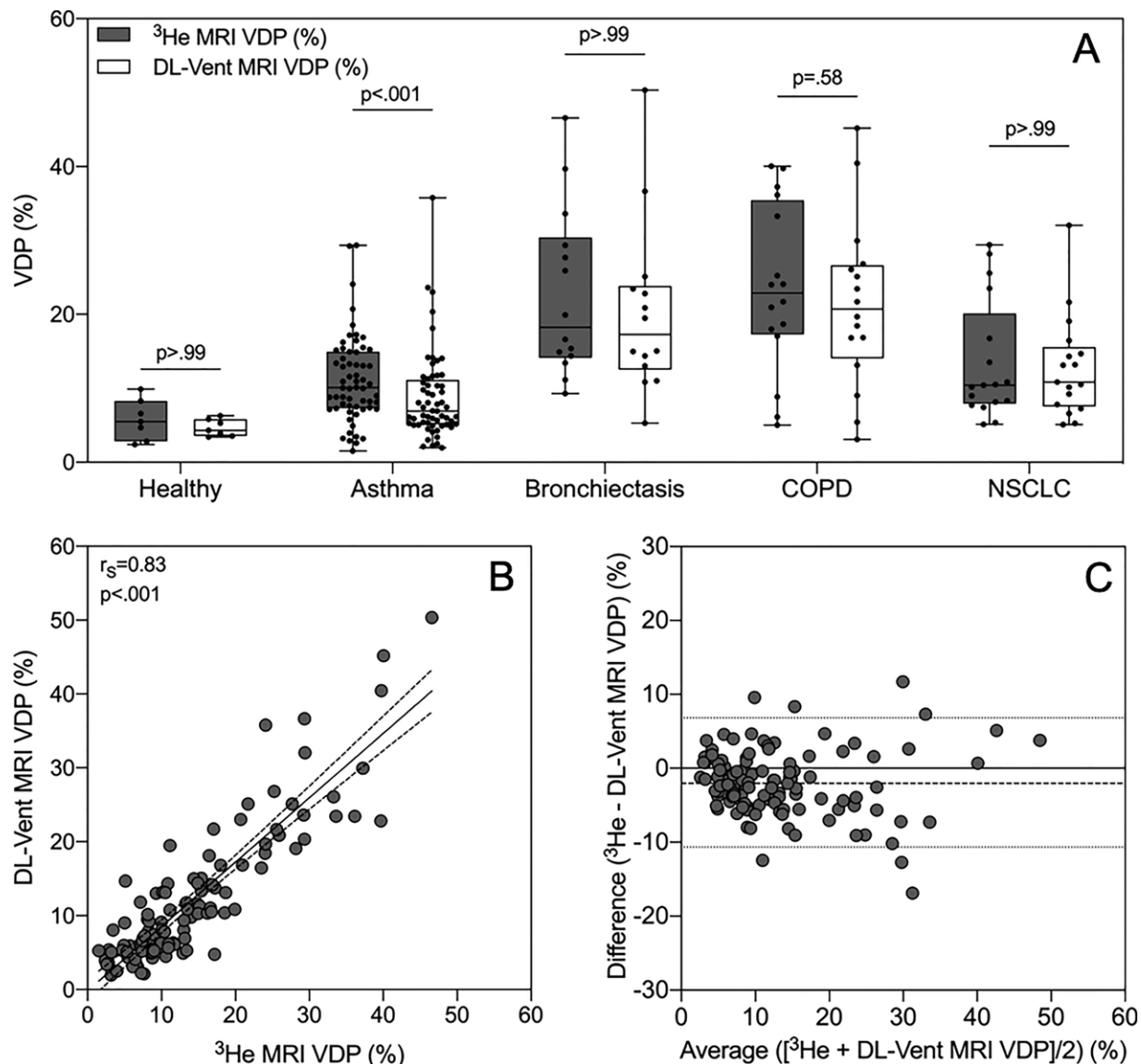


Figure 4: Comparison of and relationships between deep learning (DL) ventilation MRI (DL-Vent MRI) with helium 3 (^3He) MRI ventilation measurements. A, Comparison between DL-Vent MRI and ^3He MRI ventilation defect percentage (VDP) across all participant groups. Box-and-whisker plots show minimum, 25th percentile, median, 75th percentile, and maximum measurements with individual data points for all participants superimposed on plot. P values were determined with Wilcoxon matched-pairs signed-rank tests and adjusted with Holm-Bonferroni correction. COPD = chronic obstructive pulmonary disease, NSCLC = non-small-cell lung cancer. B, Graph shows correlation between DL-Vent MRI VDP and ^3He MRI VDP ($r_s = 0.83$, slope = 0.87, y-intercept = -0.18; $P < .001$). C, Bland-Altman analysis of agreement for DL-Vent MRI VDP with ^3He MRI VDP (mean bias = -2.0 ± 5 , lower limit = -11% , upper limit = 7%). Dotted lines indicate bias, and solid lines represent 95% CIs.

biomechanical models of dynamic structural imaging to extract functional information (33). Previous work demonstrated the use of free-breathing proton MRI to help produce ventilation-weighted maps that were validated with contrast-enhanced imaging methods (14–16). These previous approaches motivated our application of DL to this problem. Thus, this method may be leveraged to estimate ventilation maps from dynamic structural MRI in participants with lung disease. This may be particularly helpful in children with chronic pulmonary disorders, such as cystic fibrosis, where mitigation of medical radiation from repeated imaging studies is needed.

The DCNN model generated ventilation images that were quantitatively similar to ^3He MRI ventilation maps obtained

in participants. These results were further strengthened by the strong quantitative relationship between DL ventilation MRI VDP and ^3He MRI VDP, with marginal bias. We also observed no differences between DL ventilation MRI VDP and ^3He MRI VDP across all participant groups, except for participants with asthma where DL ventilation MRI VDP was greater than ^3He MRI VDP. One explanation may be that ^3He MRI ventilation defects represent the regions of the lung that high-contrast inhaled gas does not travel to and resides within a 8–15-second scan. However, defects identified by free-breathing proton MRI have been hypothesized as derived from long time-filling constants, which have previously been observed in COPD (15). Both proton density and time-filling constants may be clinically

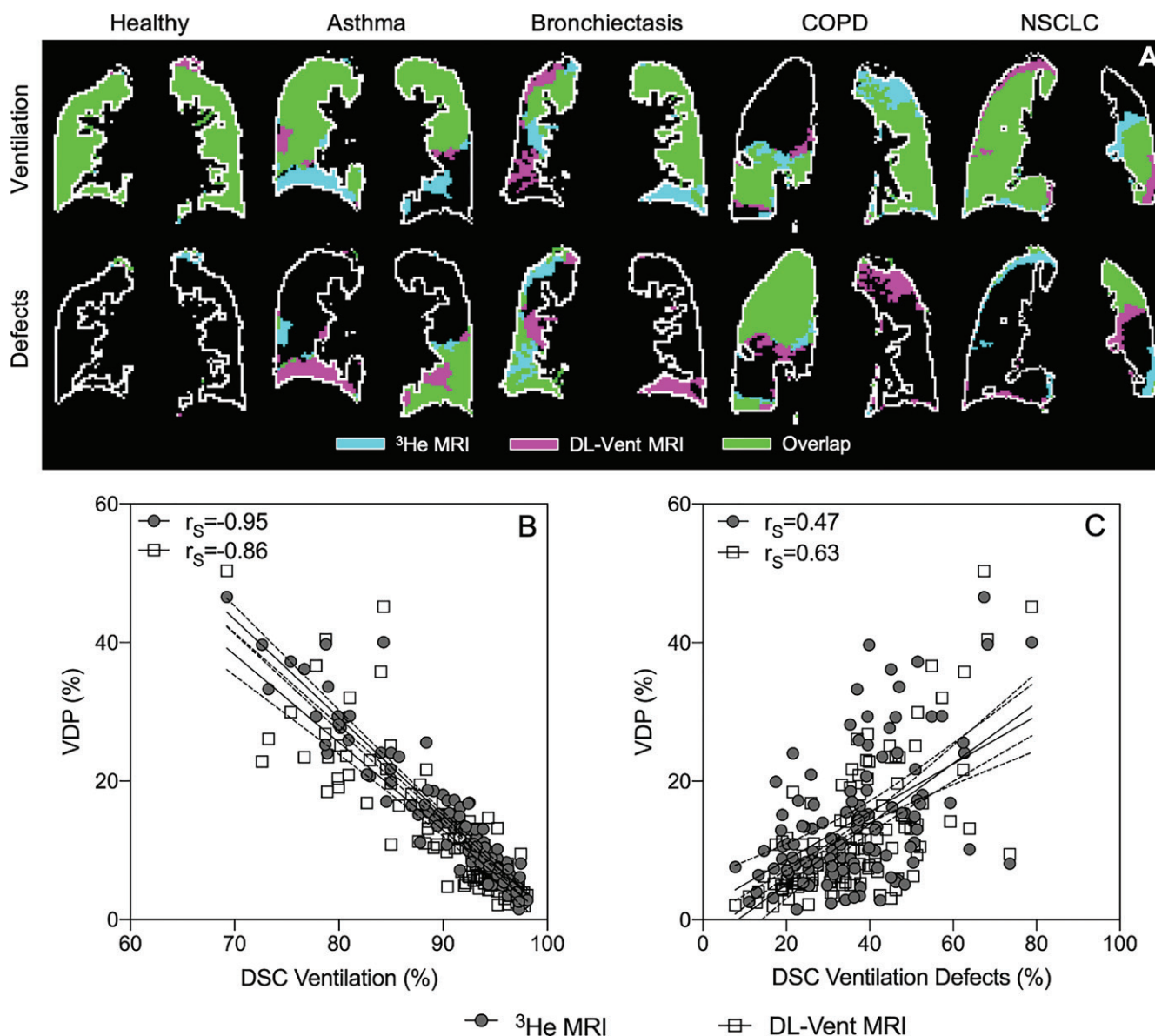


Figure 5: Spatial relationships of deep learning (DL) ventilation MRI (DL-Vent MRI) with helium 3 (^3He) MRI ventilation. A, Images show spatial overlap (green) of DL-Vent MRI (magenta) and ^3He MRI (cyan) ventilation and defects. Dice similarity coefficients (DSCs) between DL-Vent MRI and ^3He MRI ventilation were 96% in healthy volunteer; 84% in participant with asthma, 78% in participant with bronchiectasis, 84% in participant with chronic obstructive pulmonary disease [COPD], and 88% in participant with non-small cell lung cancer [NSCLC], 88%. DSCs between ventilation defects were 38% in healthy volunteer, 63% in participant with asthma, 55% in participant with bronchiectasis, 79% in participant with COPD, and 62% in participant with NSCLC. B, Graph shows correlation of DSCs between DL-Vent MRI and ^3He MRI ventilation with ^3He MRI ventilation defect percentage (VDP) ($r_s = -0.95$, $P < .001$) and DL-Vent MRI VDP ($r_s = -0.86$, $P < .001$). C, Graph shows correlation of DSCs between DL-Vent MRI and ^3He MRI between ventilation defects with ^3He MRI VDP ($r_s = 0.47$, $P < .001$) and DL-Vent MRI VDP ($r_s = 0.63$, $P < .001$). Dotted lines indicate bias, and solid lines represent 95% CIs.

Table 2: Quantitative Spatial Relationships for Ventilation and Ventilation Defects

	All (<i>n</i> = 114)	Healthy (<i>n</i> = 7)	Asthma (<i>n</i> = 60)	Bronchiectasis (<i>n</i> = 14)	COPD (<i>n</i> = 16)	NSCLC (<i>n</i> = 17)
Dice Similarity Coefficient						
Ventilation: ^3He MRI and deep learning ventilation MRI (%)	90 ± 6	97 ± 1	93 ± 4	84 ± 8	84 ± 7	91 ± 5
Ventilation Defects: ^3He MRI and deep learning ventilation MRI (%)	37 ± 14	38 ± 9	32 ± 13	41 ± 13	44 ± 15	40 ± 14

Note.—Data are means ± standard deviations. COPD = chronic obstructive pulmonary disease, ^3He = helium 3, NSCLC = non-small-cell lung cancer.

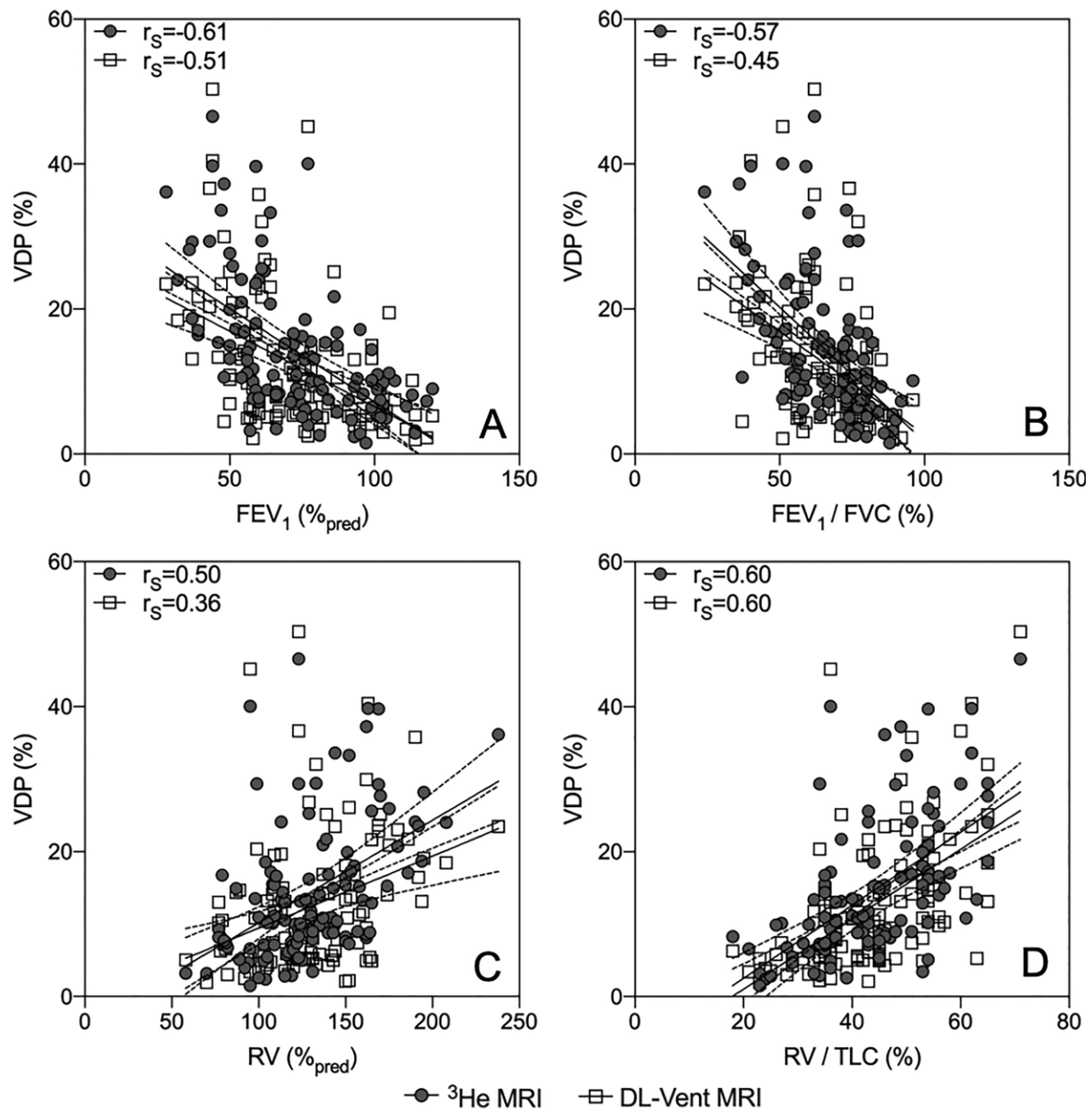


Figure 6: Graphs show relationships between helium 3 (^3He) MRI and deep learning (DL) ventilation MRI (DL-Vent MRI) ventilation measurements and pulmonary function tests. A, Forced expiratory volume in 1 second (FEV_1) linear regression with ^3He MRI ventilation defect percentage (VDP) ($r_s = -0.61$, $P < .001$) and DL-Vent MRI VDP ($r_s = -0.51$, $P < .001$). B, FEV_1 /forced vital capacity linear regression with ^3He MRI VDP ($r_s = -0.57$, $P < .001$) and DL-Vent MRI VDP ($r_s = -0.45$, $P < .001$). C, Residual volume linear regression with ^3He MRI VDP ($r_s = 0.50$, $P < .001$) and DL-Vent MRI VDP ($r_s = 0.36$, $P < .001$). D, Residual volume/total lung capacity linear regression with ^3He MRI VDP ($r_s = 0.60$, $P < .001$) and DL-Vent MRI VDP ($r_s = 0.60$, $P < .001$). Dotted lines indicate 95% CIs.

or physiologically relevant and may contribute to the overall VDP. The similar and significant relationships between DL ventilation MRI VDP and ^3He MRI VDP with pulmonary function tests also provide support that this DL model could synthesize clinically relevant ventilation images. This is an important finding that further supports and suggests that both methods provide similar functional information across a wide range of respiratory diseases.

Although DL can be used to determine functional information from anatomic information, we observed differences between DL ventilation MRI and ^3He MRI. These differences are illustrated by the DSC of ventilation defects between ^3He

and DL ventilation MRI. We should note that we observed a relationship between DSC of ventilation defects and VDP. As VDP increases, DSC of ventilation defects also increases. This is shown in Figure 5 and illustrated for the participant with COPD, where there is a large upper right lobe ventilation defect with significant overlap (DSC of ventilation defect = 77%). These observations are similar to what has previously been observed when comparing parametric response map and Fourier decomposition MRI with ^3He MRI (6,15).

DL approaches hold promise in various medical imaging applications, and herein we estimated pulmonary ventilation maps from unenhanced imaging to visualize regional ventilation

Table 3: Correlation Coefficients for DL Ventilation MRI and ^3He MRI Ventilation Defect Percentage with Pulmonary Function Measurements

Parameter	All (<i>n</i> = 114)	Healthy (<i>n</i> = 7)	Asthma (<i>n</i> = 60)	Bronchiectasis (<i>n</i> = 14)	COPD (<i>n</i> = 16)	NSCLC (<i>n</i> = 17)
Spearman correlations for ^3He MRI VDP						
DL ventilation MRI VDP	0.83	0.53	0.73	0.73	0.90	0.71
FEV ₁ (% _{pred})	−0.61	0.71	−0.50	−0.85	−0.33	−0.53
FEV ₁ /FVC (%)	−0.57*	0.27	−0.57†	−0.59	−0.54	−0.39
RV (% _{pred})	0.50*	0.00	0.44†	0.64	0.01	0.58
RV/TLC (%)	0.60*	−0.16	0.48†	0.68	0.10	0.34
Spearman correlations for DL Ventilation MRI VDP						
^3He MRI VDP	0.83	0.53	0.73	0.73	0.90	0.71
FEV ₁ (% _{pred})	−0.51	−0.14	−0.42†	−0.57	−0.18	−0.60
FEV ₁ /FVC (%)	−0.45*	0.29	−0.42†	−0.37	−0.46	−0.19
RV (% _{pred})	0.36*	−0.21	0.34†	0.33	−0.07	0.45
RV/TLC (%)	0.60*	0.09	0.42†	0.46	0.10	0.60
<i>P</i> value						
FEV ₁ (% _{pred})	.28	.15	.60	.15	.68	.78
FEV ₁ /FVC (%)	.23	.91	.30	.50	.83	.56
RV (% _{pred})	.20	.76	.54	.34	.84	.64
RV/TLC (%)	.99	.72	.70	.44	.99	.37

Note.—Spearman correlation coefficients (r_s) are for nonparametric data. The significant difference between Spearman correlations ($P < .05$) was determined with the Fisher z transformation. COPD = chronic obstructive pulmonary disease, DL = deep learning, FEV₁ = forced expiratory volume in 1 second, FVC = forced vital capacity, ^3He = helium 3, %_{pred} = percentage of predicted value, NSCLC = non-small cell lung cancer, RV = residual volume, TLC = total lung capacity, VDP = ventilation defect percentage.

* *n* = 111 participants.

† *n* = 57 participants.

abnormalities, similar to inhaled hyperpolarized gas MRI. This opens up the opportunity for functional lung imaging with conventional MRI scanners that are more universally available. Alternative indirect approaches for generating pulmonary ventilation, such as Fourier decomposition MRI (14,15), four-dimensional MRI (16), and four-dimensional CT (13), rely on complex image processing components with relatively high variability (17). This finding has prompted an international Grand Challenge (<https://www.aapm.org/GrandChallenge/CTVIE/>) specifically designed to evaluate the robustness of ventilation imaging algorithms across a range of pulmonary pathologies. DL, although it has different limitations, provides an opportunity to reduce variability and provide results similar to contrast-based ventilation imaging methods.

We must acknowledge several study limitations. First, we had a small data set size that mostly consisted of participants with asthma, so generalizability of this DL model may be somewhat limited. Second, the free-breathing MRI scans were only acquired for a singular 2D coronal section over time, limiting the spatial information. Alternative methods involving whole-lung free-breathing proton MRI would provide more spatial information, albeit a limited number of research centers are investigating both whole-lung free-breathing proton MRI and noble gas MRI scans. Third, DL models are sensitive to the data that they are trained on and typically do not scale to images obtained with different imaging protocols, field strengths, and disease phenotypes, limiting the applicability of this technique. In this study,

we simplified the input of the network to a single specific ventilation map. Although a previous study has demonstrated that specific ventilation is related to ^3He MRI ventilation (16), these two different imaging methods do not measure the same quantity, and they likely interrogate different lung function information related to airway caliber that influence the time constants for lung filling and emptying. Accordingly, in this proof-of-concept study, we developed a DCNN to map the complex relationship between free-breathing ^1H MRI-derived specific ventilation and ^3He MRI lung function that alternative approaches of measuring regional ventilation are unable to capture. Future work will involve using the entire three-dimensional free-breathing ^1H MRI time series to reduce preprocessing; the computational resources to train and test the network will need to increase as we obtain 500 free-breathing images. Fourth, the final DL ventilation MRI scans were generated by the corresponding six (different) optimal DCNNs for each fold. Consequently, a single final DCNN was not obtained; however, all six models performed relatively equivalent, as demonstrated by the performance metrics (ie, Pearson correlation coefficient, relative MAE, and relative mean error) shown in Figure 2. Finally, although DL has recently been highly successful in multiple medical imaging applications, concerns exist regarding the “black box” (34). Furthermore, data-driven convolutional neural networks may be sensitive to training data and a convolutional neural network trained on one data set may not generalize well to a different data set. Although this is a particularly challenging problem, potential solutions exist,

such as domain adaption for transferring learning and data augmentation. Future work will focus on integrating domain adaption and data augmentation techniques. Perhaps the major issue with our work was the lack of an external test set to confirm these findings.

In conclusion, in participants with diverse pulmonary pathologic findings, deep learning of free-breathing proton MRI generated pulmonary ventilation maps that showed a strong correlation with noble gas MRI and pulmonary function measurements. Our approach mapped the complex relationship between structural free-breathing proton MRI and functional inhaled hyperpolarized gas ventilation MRI. This may enable broader application of conventional free-breathing proton MRI for lung disease imaging tasks.

Author contributions: Guarantor of integrity of entire study, G.P.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, D.P.I.C., F.G., G.P.; clinical studies, D.P.I.C., G.P.; experimental studies, D.P.I.C.; statistical analysis, D.P.I.C., F.G., G.P.; and manuscript editing, D.P.I.C., F.G., G.P.

Disclosures of Conflicts of Interest: D.P.I.C. Activities related to the present article: received a grant from Natural Sciences and Engineering Research Council of Canada. Activities not related to the present article: disclosed no relevant relationships. Other relationships: disclosed no relevant relationships. F.G. disclosed no relevant relationships. L.X. disclosed no relevant relationships. G.P. disclosed no relevant relationships.

References

- Jarenbäck L, Ankerst J, Bjerner L, Tufvesson E. Acinar ventilation heterogeneity in COPD relates to diffusion capacity, resistance and reactance. *Respir Med* 2016;110:28–33.
- Lui JK, Lutchen KR. The role of heterogeneity in asthma: a structure-to-function perspective. *Clin Transl Med* 2017;6(1):29.
- Stahr CS, Samarage CR, Donnelley M, et al. Quantification of heterogeneity in lung disease with image-based pulmonary function testing. *Sci Rep* 2016;6(1):29438.
- Canadian Thoracic Society. The Human and Economic Burden of COPD: A Leading Cause of Hospital Admission in Canada. Ottawa, Canada: Canadian Thoracic Society, 2010.
- Svenningsen S, Haider E, Boylan C, et al. CT and functional MRI to evaluate airway mucus in severe asthma. *Chest* 2019;155(6):1178–1189.
- Capaldi DPI, Zha N, Guo F, et al. Pulmonary Imaging Biomarkers of Gas Trapping and Emphysema in COPD: (3)He MR Imaging and CT Parametric Response Maps. *Radiology* 2016;279(2):597–608.
- Nakano Y, Müller NL, King GG, et al. Quantitative assessment of airway remodeling using high-resolution CT. *Chest* 2002;122(6,Suppl):271S–275S.
- Hayhurst MD, MacNee W, Flenley DC, et al. Diagnosis of pulmonary emphysema by computerised tomography. *Lancet* 1984;2(8398):320–322.
- Kauczor HU, Hofmann D, Kreitner KF, et al. Normal and abnormal pulmonary ventilation: visualization at hyperpolarized He-3 MR imaging. *Radiology* 1996;201(2):564–568.
- Park EA, Goo JM, Park SJ, et al. Chronic obstructive pulmonary disease: quantitative and visual ventilation pattern analysis at xenon ventilation CT performed by using a dual-energy technique. *Radiology* 2010;256(3):985–997.
- King GG, Eberl S, Salome CM, Meikle SR, Woolcock AJ. Airway closure measured by a technegas bolus and SPECT. *Am J Respir Crit Care Med* 1997;155(2):682–688.
- Brudin LH, Rhodes CG, Valind SO, Buckingham PD, Jones T, Hughes JM. Regional structure-function correlations in chronic obstructive lung disease measured with positron emission tomography. *Thorax* 1992;47(11):914–921.
- Guerrero T, Sanders K, Noyola-Martinez J, et al. Quantification of regional ventilation from treatment planning CT. *Int J Radiat Oncol Biol Phys* 2005;62(3):630–634.
- Bauman G, Puderbach M, Deimling M, et al. Non-contrast-enhanced perfusion and ventilation assessment of the human lung by means of fourier decomposition in proton MRI. *Magn Reson Med* 2009;62(3):656–664.
- Capaldi DPI, Sheikh K, Guo F, et al. Free-breathing pulmonary 1H and Hyperpolarized 3He MRI: comparison in COPD and bronchiectasis. *Acad Radiol* 2015;22(3):320–329.
- Capaldi DPI, Eddy RL, Svenningsen S, et al. Free-breathing Pulmonary MR Imaging to Quantify Regional Ventilation. *Radiology* 2018;287(2):693–704.
- Hegi-Johnson F, de Ruyscher D, Keall P, et al. Imaging of regional ventilation: Is CT ventilation imaging the answer? A systematic review of the validation data. *Radiother Oncol* 2019;137:175–185.
- Tahir BA, Hughes PJC, Robinson SD, et al. Spatial comparison of CT-based surrogates of lung ventilation with hyperpolarized Helium-3 and Xenon-129 gas MRI in patients undergoing radiation therapy. *Int J Radiat Oncol Biol Phys* 2018;102(4):1276–1286.
- Kipritidis J, Tahir BA, Cazoulat G, et al. The VAMPIRE challenge: A multi-institutional validation study of CT ventilation imaging. *Med Phys* 2019;46(3):1198–1217.
- Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Med Image Anal* 2017;42:60–88.
- Han X. MR-based synthetic CT generation using a deep convolutional neural network method. *Med Phys* 2017;44(4):1408–1419.
- Capaldi DPI, Sheikh K, Eddy RL, et al. Free-breathing Functional Pulmonary MRI: Response to Bronchodilator and Bronchoprovocation in Severe Asthma. *Acad Radiol* 2017;24(10):1268–1276.
- Miller MR, Hankinson J, Brusasco V, et al. Standardisation of spirometry. *Eur Respir J* 2005;26(2):319–338.
- Wanger J, Clausen JL, Coates A, et al. Standardisation of the measurement of lung volumes. *Eur Respir J* 2005;26(3):511–522.
- Kirby M, Heydarian M, Svenningsen S, et al. Hyperpolarized 3He magnetic resonance functional imaging semiautomated segmentation. *Acad Radiol* 2012;19(2):141–152.
- Heinrich MP, Jenkinson M, Bhushan M, et al. MIND: modality independent neighbourhood descriptor for multi-modal deformable registration. *Med Image Anal* 2012;16(7):1423–1435.
- Guo F, Svenningsen S, Eddy RL, et al. Anatomical pulmonary magnetic resonance imaging segmentation for regional structure-function measurements of asthma. *Med Phys* 2016;43(6):2911–2926.
- Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab N, Hornegger J, Wells W, Frangi A, eds. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. MICCAI 2015. Lecture Notes in Computer Science, vol 9351. Cham, Switzerland: Springer, 2015; 234–241.
- LeCun YA, Bottou L, Orr GB, Müller KR. Efficient BackProp. In: Montavon G, Orr GB, Müller KR, et al, eds. *Neural Networks: Tricks of the Trade*. Lecture Notes in Computer Science, vol 7700. Berlin, Germany: Springer, 2012; 9–48.
- Dice LR. Measures of the amount of ecologic association between species. *Ecology* 1945;26(3):297–302.
- Kirby M, Svenningsen S, Owringi A, et al. Hyperpolarized 3He and 129Xe MR imaging in healthy volunteers and patients with chronic obstructive pulmonary disease. *Radiology* 2012;265(2):600–610.
- Altman DG. *Practical statistics for medical research*. New York, NY: Chapman & Hall/CRC, 1990.
- Young HM, Eddy RL, Parraga G. MRI and CT lung biomarkers: Towards an in vivo understanding of lung biomechanics. *Clin Biomech (Bristol, Avon)* 2019;66:107–122.
- Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat Mach Intell* 2019;1(5):206–215.