A Learning Framework for Predicting CT-based PRM Biomarker from MRI Sequences in COPD

Yiling Xu[1,2], Simon M. F. Triphan[1,2,3], Julian Grolig[1,2], Hanyi Zhang[1], Jürgen Biederer[1,2,3], Craig J. Galbán[5,6], Hans-Ulrich Kauczor[1,2,3], Mark O. Wielpütz[2,4], and Oliver Weinheimer[1,2,3⋆]

[1] Department of Diagnostic and Interventional Radiology, University Hospital of Heidelberg, Heidelberg, Germany
weinheimer@uni-heidelberg.de
[2] Translational Lung Research Center Heidelberg (TLRC), German Center for Lung Research (DZL), University of Heidelberg, Heidelberg, Germany
[3] Department of Diagnostic and Interventional Radiology with Nuclear Medicine, Thoraxklinik at University of Heidelberg, Heidelberg, Germany
[4] Department of Diagnostic Radiology and Neuroradiology, University Medicine Greifswald, Greifswald, Germany
[5] Department of Biomedical Engineering, University of Michigan, Ann Arbor, United States
[6] Department of Radiology, University of Michigan, Ann Arbor, United States

Abstract. Image-based biomarkers provide non-invasive regional assessment of structural-functional abnormalities in Chronic Obstructive Pulmonary Disease (COPD). For example, quantitative computed tomography (QCT) identifies emphysema and small airway disease, while functional MRI measures lung ventilation and perfusion. In recent years, machine learning techniques have been introduced to predict quantitative indices from alternative imaging modalities, with the aim to reduce scanning time, radiation dose and/or costs in the clinical setting. However, most of those works focused on lung ventilation, while robust quantification of regional lung perfusion of dynamic contrast-enhanced (DCE) MRI remains a challenging task. In addition, previous studies focused only on learning from a single imaging modality. In this study, we explore a deep learning-based model to predict conventionally CT-based biomarkers, namely Parametric response mapping (PRM) classifications, from multi-sequence structural-functional MR images. Our proposed model achieves very strong correlations in predicting $\%\mathrm{PRM}_{\mathrm{emphysema}}$ (Pearson correlation coefficient $r = 0.91, \quad p < 0.001$ at patient level and $r = 0.87, \quad p < 0.001$ at lung lobe level), and moderate to strong correlations in predicting $\%\mathrm{PRM}_{\mathrm{normal}}$ ($r = 0.60, \quad p < 0.001$ at patient level and $r = 0.58, \quad p < 0.001$ at lung lobe level) in unseen COPD patients.

Keywords: Functional Lung Imaging · Multi-modal Medical Imaging · COPD

⋆ Corresponding author

1 Introduction

Imaging techniques are of increasing importance in the assessment of COPD, especially for regional pulmonary abnormalities. The Parametric response mapping (PRM) method introduced by Galbán et al. [5] spatially aligns the expiratory to the inspiratory CT scans by a deformable volumetric registration process. Thereafter, the classification of lung voxels can be carried out by using the expiratory and inspiratory attenuation value for every voxel. This allows classifying lung voxels as normal parenchyma (PRM_{normal}), functional small-airway disease (PRM_{fSAD}), and emphysema ($PRM_{emphysema}$). As an emerging technique in pulmonary imaging, MRI provides a radiation-free alternative and numerous possibilities in functional imaging, such as ventilation and perfusion within the lungs.

Attempts have been made to combine imaging-based indices from different modalities. MacNeil et al. [11] proposed a comprehensive biomarker that combines CT with hyperpolarized helium MRI. The emergence of deep learning facilitates the prediction of lung function parameters across different imaging modalities. For example, ventilation maps were synthesized from free-breathing proton MRI. [4], or from paired inspiratory-expiratory non-contrast CT [1].

In the meantime, such learning-based methods remain underexplored toward functional perfusion MRI. Quantitative perfusion indices such as pulmonary blood flow (PBF) and pulmonary blood volume (PBV) can be drawn from dynamic contrast-enhanced MRI (DCE-MRI). However, such indices suffer from inherent problems such as low signal-to-noise ratio, nonlinearity of the contrast agent, and motion artifacts.

To address those problems, Schiwek et al. [13] proposed a threshold-based algorithm too discern well-perfused and perfusion-defect regions, which showed moderate to strong correlations to %PRM indices and lung function tests. Further studies [14] suggest that certain T1 signals are related to lung perfusion.

Moreover, most of the published studies focus on utilizing source images from a single modality or a single MR sequence. We believe a more comprehensive model that takes versatile structural-functional images as input can learn additive features from different measurements, alleviate the limitations of the low Signal-to-Noise Ratio (SNR) and low spatial resolution in individual measurements, thus producing better results. Such a model may be employed for more challenging tasks as to predict conventionally CT-based biomarkers from MRI. Inspired by multi-sequence MRI synthesis [6], we proposed a model that leverages complementary information from multi-sequence structural-functional MRI to predict voxel-wise PRM classifications. The model can be alternatively applied in the scenario of missing sequences.

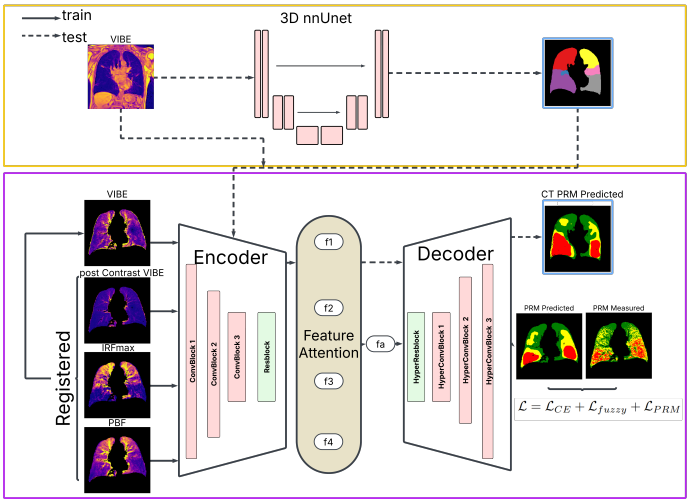

Fig. 1. Overview of the proposed framework. The upper part is an nnUnet trained to segment lung lobes on VIBE MRI. The lower part composes an encoder-decoder-based image translation model which predicts CT-driven PRM maps from structural and quantitative functional MRI. The two models were trained separately but with the same data split to keep the test set unseen. In the test phase, lung segmentation and lobe segmentation from the trained nnUnet were used as lung mask and forwarded to the PRM prediction model, to determin predicted $\%\text{PRM}_{\text{emphysema}}$, $\%\text{PRM}_{\text{fSAD}}$, and $\%\text{PRM}_{\text{normal}}$ both on the patient-level and lobe-level.

2 Materials and Methods

2.1 Data Set

104 patients with COPD (age 56.9±18.6, mean GOLD 2.10±1.19) were enrolled at three different imaging centers as a subset of the COSYCONET multi-center trial [8]. All patients underwent same-day paired inspiratory CT and structural-functional MR scans, including a 4D dynamic contrast enhanced (DCE) MRI.

MRI examinations were performed using identical models of 1.5 T MR scanner (Magnetom Aera, Siemens Healthineers) and a standardized chest protocol. For DCE-MRI, a time-resolved angiography with interleaved stochastic trajectories (TWIST) sequence with gadolinium-based contrast agent followed by a saline chaser was acquired, with a slice thickness of 5.0 mm, and a coronal-plane resolution between 1.57 × 1.57 mm to 2.34 × 2.34 mm. Time-resolved residue function maps (R(t)-map) were computed, and maximum contrast enhancement (IRmax) map and PBF map were quantified accordingly with established methods [12]. Furthermore, this study included T1-weighted 3D-GRE structural Volumetric Interpolated Breath-hold Examination (VIBE) measurements obtained both before and after DCE-MRI (post VIBE), with a slice thickness of 4.0 mm,

and a coronal-plane resolution between 0.98×0.98 mm to 1.46×1.46 mm. MR images used in the study were all acquired at inspiratory breath hold.

Non-contrast low-dose CTs (from three different Siemens Somatom scanners) with paired scans in inspiratory and expiratory breath-hold were acquired. Images were reconstructed with a soft kernel, with a slice thickness of 1.0 mm, and a in-plane resolution between 0.49×0.49 mm to 0.94×0.94 mm. Lung segmentations, lung lobe segmentations, and voxel-wise PRM classifications were generated with an automated in-house software [3, 15, 10, 9]. We further define $PRM_{\text{abnormal}} = PRM_{\text{emphysema}} \cup PRM_{\text{fSAD}}$.

All images were co-registered to the spatial layout of VIBE MRI. TWIST MRI (used to compute PBF and IRFmax) and post-contrast VIBE images were registered to morphological VIBE MRI using affine and deformable registration, with aim of correcting minor positioning or breathing state shifts. CT images, as acquired at higher resolution and potentially different field-of-view, were first resampled to match the spacing of VIBE MRI, and subsequently aligned using ANTspy SyN framework [2] which consists of rigid, affine and deformable registration. The computed spatial transformation was hence applied to lobe segmentation and PRM maps as well. A registered CT lung mask was applied before images were normalized to [-1,1].

We randomly split samples into 60 for training, 14 for validation and 30 for testing. All images were cropped or padded to the size of $60 \times 256 \times 256$ before being forwarded to the each of the following models.

2.2 The Learning Framework

The proposed learning framework is composed of two parts: A lung lobe segmentation model to identify six lung lobes on non-contrast VIBE MRI, and a PRM prediction model to classify each lung voxel from multi-sequence structural functional MRI. An overall architecture is presented in Fig. 1.

Lung Lobe Segmentation Model For training the model, lung lobes were segmented from inspiratory CT using the CT analysis software ***** [10]. Lung halves were identified and each were further segmented in to upper, middle, and lower lobes, resulting in six lung lobes for each image, except for two cases where the lingula (left middle lobe) could not be identified. Results were then spatially registered to VIBE MRI as ground truth. We employed an nnUnet [7] to identify lung lobes in the non-contrast VIBE MRI. The model was self-configured and trained for 1000 epochs.

Multi-sequence PRM Prediction Model For prediction of a voxel-wise PRM classification, we adapted the multi-sequence MR synthesis model [6], which consists of an image encoder, a feature attention layer and an image decoder. Modifications were made as larger convolution kernels of $5 \times 5 \times 5$ were used for all ConvBlocks. Channels numbers were set as 8, 16, and 32 for the encoding path, and 32, 8, 3 for the decoding path. Co-registered 3D MR images

and quantitative maps were forwarded. The model was trained for 200 epochs with a learning rate of 1×10^{-4} and a batch size of 8. Early stopping was employed and the model with best validation performance data was retained for the evaluation on test set.

To minimize the influence of the partial volume effect tangled with potential bias introduced by image registration, we implemented a loss function that comprises three parts: a cross-entropy loss of the voxel-wise PRM classification ($\mathcal{L}_{\text{CE}}$), a fuzzy classification loss ($\mathcal{L}_{\text{fuzzy}}$) where the labels were first blurred with a gaussian filter with a size of $3 \times 9 \times 9$, and a total lung PRM percentage loss ($\mathcal{L}_{\text{PRM}}$). In our experimental setting, the loss weights λ_1, λ_2, and λ_3 were set as 0.3, 0.3, and 0.4 respectively.

$$\mathcal{L} = \lambda_1 \mathcal{L}_{\text{CE}} + \lambda_2 \mathcal{L}_{\text{fuzzy}} + \lambda_3 \mathcal{L}_{\text{PRM}}$$

where

$$\mathcal{L}_{\text{CE}} = -\frac{1}{N} \sum_{i=1}^{N} \sum_{c \in \{\text{emph,fSAD,norm}\}} GT_i^{(c)} \log \left(p_i^{(c)} \right)$$

$$\mathcal{L}_{\text{fuzzy}} = -\frac{1}{N} \sum_{i=1}^{N} \sum_{c \in \{\text{emph,fSAD,norm}\}} blurred(GT^{(c)})_i \log \left(p_i^{(c)} \right)$$

$$\mathcal{L}_{\text{PRM}} = \frac{1}{3} \sum_{c \in \{\text{emph,fSAD,norm}\}} |(\%p_c - \%GT_c|$$

3 Results

The proposed lobe segmentation model was evaluated against the CT lobe segmentation registered to VIBE MRI, achieving a DICE score of 0.97 ± 0.01 for the total lung segmentation, and 0.850 ± 0.06 for the lobe segmentation.

The lung masks generated were subsequently applied to the test set and forwared to the trained PRM prediction model. Fig. 2) shows representative slices of the predicted PRM classifications in 3 subjects at different COPD GOLD stages. The proposed model achieves an overall DICE score of 0.69 ± 0.18 for identifying $\text{PRM}_{\text{abnormal}}$, and of 0.45 ± 0.22 for identifying $\text{PRM}_{\text{emphysema}}$ among the 30 patients.

Results of the two models were further integrated to compute patient-level and lobe-level predicted $\%\text{PRM}_{\text{emphysema}}$, $\%\text{PRM}_{\text{fSAD}}$, and $\%\text{PRM}_{\text{normal}}$. For comparison, we computed the measured CT-driven PRM directly from the original scan to avoid possible biases induced by image registration.

We calculated the Pearson correlation coefficient for the comparison of predicted and measured %PRM. The results in Fig. 3 show that our model achieves a very strong correlation in predicting $\%\text{PRM}_{\text{emphysema}}$ ($r = 0.91, \quad p < 0.001$ at patient level and $r = 0.87, \quad p < 0.001$ at lobe level), a weak correlation in predicting $\%\text{PRM}_{\text{fSAD}}$ ($r = 0.15, \quad p = 0.42$ at patient level and $r =$

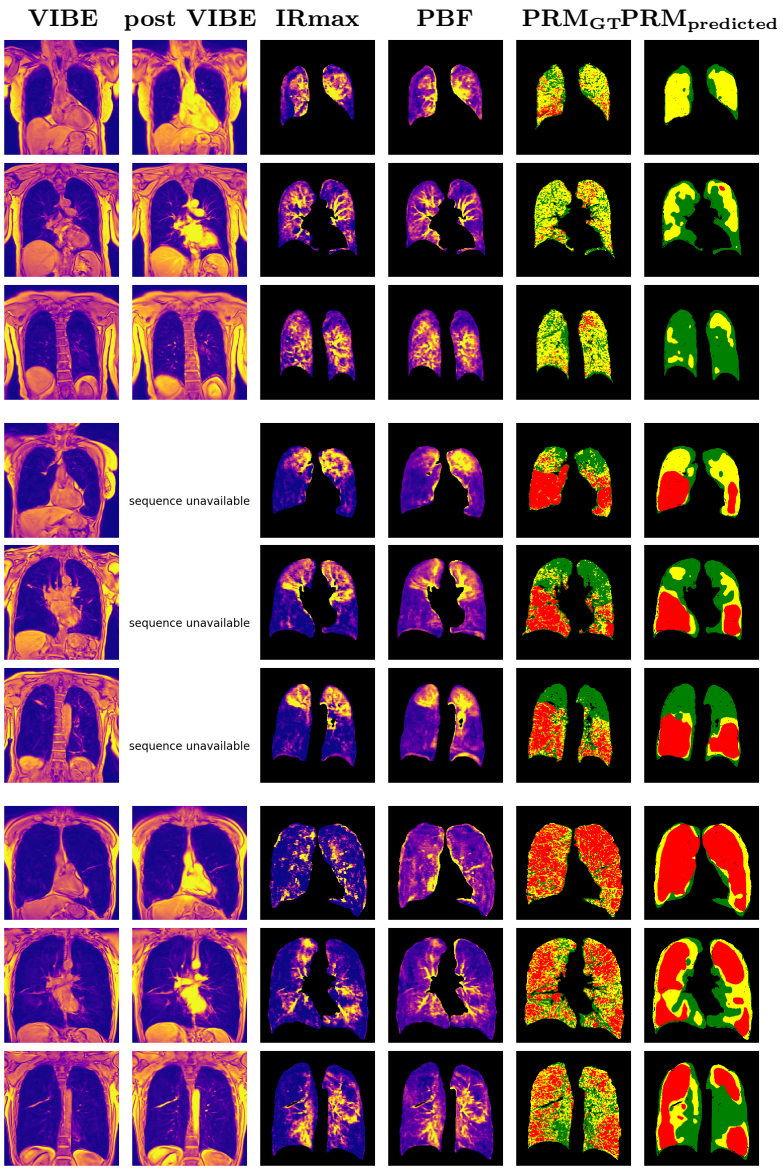

Fig. 2. Representative results from the unseen test set. Row 1-3 from a subject with GOLD 1 ($DICE_{emph}$=0.36, $DICE_{abnormal}$=0.796). Row 4-6 from a subject with GOLD3 ($DICE_{emph}$=0.70, $DICE_{abnormal}$=0.799, in this case post VIBE sequence could not be reconstructed, the model was tested on the other 3 input images). ROW 7-9 from a subject with GOLD 4 ($DICE_{emph}$=0.62, $DICE_{abnormal}$=0.842). VIBE, IRMax and PBF images are shown in relative intensity, PRM mpas were color coded as such: red represents $PRM_{emphysema}$, yellow represents PRM_{fSAD}, and gree represents PRM_{normal}.

0.32, $p < 0.001$ at lobe level) and a moderate to strong correlation in predicting $\%\text{PRM}_{\text{normal}}$ ($r = 0.60$, $p < 0.001$ at patient level and $r = 0.58$, $p < 0.001$ at lobe level). The mean differences are $-1.8\% \pm 9.0\%$ for $\text{PRM}_{\text{emphysema}}$, $-4.6\% \pm 21.0\%$ for PRM_{fSAD}, and $-0.1\% \pm 20.8\%$ for $\text{PRM}_{\text{normal}}$, respectively.

We further assessed the model's performance in the scenario with missing input sequences. $\text{DICE}_{\text{emph}}$, $\text{DICE}_{\text{abnormal}}$, and the correlation coefficient against the measured %PRM were computed on the prediction when different configurations of input images were forwarded to the trained model.

As the results shown in Table 1, the highest overall performance is obtained when the complete set of four source images are provided. Notably, with the exclusion of the post VIBE measurement alone, the model still yields comparable results, while removing VIBE measurement induces minor performance degradation. In contrast, omitting functional images, IRmax or PBF, leads to pronounced impairment of the model capacity, which highlights the contribution of functional maps in the PRM prediction task.

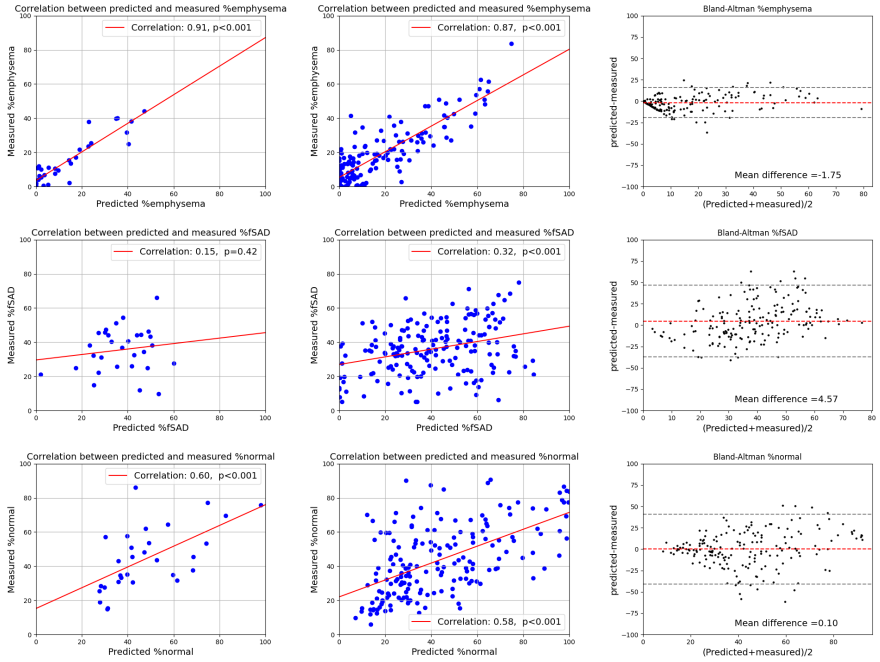

Fig. 3. Comparison between predicted $\%\text{PRM}_{\text{emphysema}}$, $\%\text{PRM}_{\text{fSAD}}$, and $\%\text{PRM}_{\text{normal}}$ against the measured values on CT. left: scatter plot with Pearson's correlation coefficient at patient level, middle: scatter plot with Pearson's correlation coefficient at lobe level, right: corresponding Bland-Altman plot, the red lines represent mean difference, and the dashed lines represent the values corresponding to $\pm 1.95\sigma$.

Table 1. Model performance with different input image perturbations, in comparison with established qMR. *represents $p < 0.05$.

Model with Input Images	$DICE_{emph}$	$DICE_{abnormal}$	$r\%_{emph}$	$r\%_{fSAD}$	$r\%_{normal}$
All 4	$\mathbf{0.45 \pm 0.22}$	0.69 ± 0.18	$\mathbf{0.91}^*$	0.14	$\mathbf{0.60}^*$
VIBE, IRmax, PBF	0.43 ± 0.22	$\mathbf{0.70 \pm 0.15}$	$\mathbf{0.88}^*$	0.23	$\mathbf{0.62}^*$
post VIBE, IRmax, PBF	$\mathbf{0.44 \pm 0.20}$	$\mathbf{0.72 \pm 0.15}$	0.86^*	0.13	0.46^*
VIBE, post VIBE, IRmax	0.21 ± 0.20	0.50 ± 0.15	0.41^*	0.28	0.13
VIBE, post VIBE, PBF	0.11 ± 0.07	0.33 ± 0.14	-0.16	0.13	-0.15
VIBE, IRmax	0.34 ± 0.20	0.49 ± 0.24	$\mathbf{0.89}^*$	$\mathbf{0.33}^*$	$\mathbf{0.70}^*$
IRFmax, PBF	$\mathbf{0.44 \pm 0.20}$	$\mathbf{0.72 \pm 0.15}$	0.85^*	-0.14	0.44^*
qMR					
QDP (derived from IRmax)[13]	–	–	0.61^*	-0.18^*	-0.6^*
mean PBF	–	–	-0.08	-0.20	0.22

4 Discussion

We proposed a deep learning-based framework that predicts CT-based regional abnormalities from multi-sequence structural functional MRI. The framework was evaluated on 104 patients at different stage of COPD, and results were compared both at patient level and lung lobe level, as well as the regional overlap.

The proposed model effectively identifies emphysema, achieving very strong correlation and small mean difference against the ground truth obtained from high resolution CT images. The multi-sequence data can compensate the low SNR and low spatial resolution of a single modality, and may provide a radiation-free alternative to determine emphysema in COPD patients.

In the meantime, defining small airway obstruction regions remains challenging, as only a weak correlation was achieved. The results align with the fact that $\%PRM_{fSAD}$ are identified from CT scans acquired at inspiration and expiration, while in this study we only utilized MR sequences at inspiratory status. Other studies [4, 1] suggest, that the additional inclusion of a scan at expiration can provide supplementary ventilation information, which may further improve the prediction of $\%PRM_{fSAD}$.

In addition, we evaluated the proposed model across varied input image combinations, thereby highlighting the critical contribution of the individual measurements, especially IRmax and PBF. In future work, we will incorporate additional sequences and assess their ability to contribute complementary information.

We acknowledge the following limitations of this study. The model was only validated on a relatively small dataset, since multi-sequence MRI data of COPD patients are not commonly available. Conversely, the flexible framework allows us to train the model on versatile input images, and the employment of quantitative images as input should enhance the generalization of the proposed model. We plan to test the model on a larger data set in future studies.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Astley, J.R., Biancardi, A.M., Marshall, H., Hughes, P.J., Collier, G.J., Hatton, M.Q., Wild, J.M., Tahir, B.A.: A hybrid model-and deep learning-based framework for functional lung image synthesis from multi-inflation ct and hyperpolarized gas mri. Medical Physics **50**(9), 5657–5670 (2023)
2. Avants, B.B., Tustison, N., Song, G., et al.: Advanced normalization tools (ants). Insight j **2**(365), 1–35 (2009)
3. Buschsieweke, C., Heussel, C., Thelen, M., Kauczor, H.: Fully automatic detection and quantification of emphysema on thin section md-ct of the chest by a new and dedicated software. RoFo: Fortschritte auf dem Gebiete der Rontgenstrahlen und der Nuklearmedizin **176**(10), 1409–1415 (2004)
4. Capaldi, D.P., Guo, F., Xing, L., Parraga, G.: Pulmonary ventilation maps generated with free-breathing proton mri and a deep convolutional neural network. Radiology **298**(2), 427–438 (2021)
5. Galbán, C.J., Han, M.K., Boes, J.L., Chughtai, K.A., Meyer, C.R., Johnson, T.D., Galbán, S., Rehemtulla, A., Kazerooni, E.A., Martinez, F.J., et al.: Computed tomography–based biomarker provides unique signature for diagnosis of copd phenotypes and disease progression. Nature medicine **18**(11), 1711–1715 (2012)
6. Han, L., Tan, T., Zhang, T., Huang, Y., Wang, X., Gao, Y., Teuwen, J., Mann, R.: Synthesis-based imaging-differentiation representation learning for multi-sequence 3d/4d mri. Medical Image Analysis **92**, 103044 (2024)
7. Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H.: nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. Nature methods **18**(2), 203–211 (2021)
8. Karch, A., Vogelmeier, C., Welte, T., Bals, R., Kauczor, H.U., Biederer, J., Heinrich, J., Schulz, H., Gläser, S., Holle, R., et al.: The german copd cohort cosyconet: aims, methods and descriptive analysis of the study population at baseline. Respiratory medicine **114**, 27–37 (2016)
9. Konietzke, P., Weinheimer, O., Triphan, S.M.F., Nauck, S., Wuennemann, F., Konietzke, M., Jobst, B.J., Jörres, R.A., Vogelmeier, C.F., Heussel, C.P., Kauczor, H.U., Wielpütz, M.O., Biederer, J.: Gold grade-specific characterization of copd in the cosyconet multi-center trial: comparison of semiquantitative mri and quantitative ct. European Radiology (1 2025). https://doi.org/10.1007/s00330-024-11269-3
10. Konietzke, P., Weinheimer, O., Wielpütz, M.O., Savage, D., Ziyeh, T., Tu, C., Newman, B., Galbán, C.J., Mall, M.A., Kauczor, H.U., et al.: Validation of automated lobe segmentation on paired inspiratory-expiratory chest ct in 8-14 year-old children with cystic fibrosis. PloS one **13**(4), e0194557 (2018)
11. MacNeil, J.L., Capaldi, D.P., Westcott, A.R., Eddy, R.L., Barker, A.L., McCormack, D.G., Kirby, M., Parraga, G.: Pulmonary imaging phenotypes of chronic obstructive pulmonary disease using multiparametric response maps. Radiology **295**(1), 227–236 (2020)
12. Ohno, Y., Hatabu, H., Murase, K., Higashino, T., Kawamitsu, H., Watanabe, H., Takenaka, D., Fujii, M., Sugimura, K.: Quantitative assessment of regional pulmonary perfusion in the entire lung using three-dimensional ultrafast dynamic

contrast-enhanced magnetic resonance imaging: Preliminary experience in 40 subjects. Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine **20**(3), 353–365 (2004)
13. Schiwek, M., Triphan, S.M., Biederer, J., Weinheimer, O., Eichinger, M., Vogelmeier, C.F., Jörres, R.A., Kauczor, H.U., Heußel, C.P., Konietzke, P., et al.: Quantification of pulmonary perfusion abnormalities using dce-mri in copd: comparison with quantitative ct and pulmonary function. European Radiology pp. 1–12 (2022)
14. Triphan, S.M., Konietzke, M., Biederer, J., Eichinger, M., Vogelmeier, C.F., Jörres, R.A., Kauczor, H.U., Heußel, C.P., Jobst, B.J., Wielpütz, M.O., et al.: Echo time-dependent observed t1 and quantitative perfusion in chronic obstructive pulmonary disease using magnetic resonance imaging. Frontiers in Medicine **10**, 1254003 (2024)
15. Weinheimer, O., Achenbach, T., Heussel, C.P., Düber, C.: Automatic lung segmentation in mdct images. In: Fourth International Workshop on Pulmonary Image Analysis. vol. 2011, pp. 241–255. CreateSpace 2011 September 18; Toronto, Canada (2011)