Published in final edited form as:

Acad Radiol. 2019 March ; 26(3): 383–394. doi:10.1016/j.acra.2018.06.024.

A model for predicting future FEV1 decline in smokers using hyperpolarized ^3He Magnetic Resonance Imaging

Ryan J. Baron¹, Hooman Hamedani¹, Stephen J. Kadlecsek¹, Ian F. Duncan¹, Yi Xin¹, Sarmad Siddiqui¹, Mehrdad Pourfathi¹, Maurizio Cereda², and Rahim R. Rizi¹

¹Department of Radiology, University of Pennsylvania, Philadelphia, PA 19104, USA

²Department of Anesthesiology and Critical Care, University of Pennsylvania, Philadelphia, PA 19104, USA

Abstract

Rationale and Objectives: The purpose of this study was to assess the effectiveness of hyperpolarized (HP) helium-3 magnetic resonance (MR)-based imaging markers in predicting future FEV1 decline/COPD progression in smokers compared to current diagnostic techniques.

Materials and Methods: 60 total subjects (15 nonsmokers and 45 smokers) participated in both baseline and follow-up visits (~1.4 years apart). At both visits, subjects completed pulmonary function testing (PFT), a six-minute walk test (6MWT), and the St. George Respiratory Questionnaire (SGRQ). Using helium-3 MR imaging, means (M) and standard deviations (H) of oxygen tension (P_{AO_2}), fractional ventilation (FV), and apparent diffusion coefficient (ADC) were calculated across 12 regions of interest (ROIs) in the lungs. Subjects who experienced FEV1 decline >100 mL/year were deemed “decliners,” while those who did not were deemed “sustainers.” Non-imaging and imaging prediction models were generated through a logistic regression model, which utilized measurements from sustainers and decliners.

Results: The non-imaging prediction model included the SGRQ total score, DLCO/VA, and distance walked in 6MWT. An ROC curve for this model yielded a sensitivity of 75% and specificity of 68% with an overall area under the curve (AUC) of 65%. The imaging prediction model generated following the same methodology included ADC^{H} , FV^{H} , and $P_{\text{AO}_2}^{\text{H}}$. The resulting ROC curve yielded a sensitivity of 87.5%, specificity of 82.8%, and an AUC of 89.7%.

Conclusion: The imaging prediction model generated from measurements obtained during ^3He MR imaging is better able to predict future FEV1 decline compared to one based on current clinical tests and demographics. The imaging model’s superiority appears to arise from its ability to distinguish well-circumscribed, severe disease from a more uniform distribution of moderately altered lung function, which is more closely associated with subsequent FEV1 decline.

Corresponding Author: Rahim R. Rizi, Hospital of the University of Philadelphia, MRI Learning Center, 3400 Spruce St, 1 Founders Pavilion, Philadelphia, PA 19104, Phone: (215) 615-2426, rahim.rizi@uphs.upenn.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

Noble Gas MRI; Helium-3 MRI; FEV₁ Decline; COPD

Introduction

Chronic obstructive pulmonary disorder (COPD) is a progressive disease of the lungs in which obstructed airflow and destruction of the lung parenchyma cause disrupted and deteriorating lung function often accompanied by sputum production, dyspnea, and cough (1). In 2015, 3.17 million deaths were attributed to COPD, accounting for nearly 5% of total global deaths (2). COPD is typically diagnosed by assessing forced expiratory volume in one second (FEV₁) together with the presentation of associated symptoms; however, while tracking annual FEV₁ decline has been considered the most effective way to monitor COPD progression since the 1970s (3–5), it has proven ineffective for predicting future lung function decline. In recent years, the hyperpolarized noble gases have been developed as MRI contrast agents capable of delivering images of gas distribution throughout all regions of the lung, which can be used to quantify lung function during ventilation and diffusion (6,7). As a result, assessing gas distribution and heterogeneity in signal intensity throughout the lung now offers an alternative method for monitoring COPD progression with better prognostic potential. In this paper, we present a model for using ³He MR imaging markers to predict functional decline in the lungs of smokers with high sensitivity and specificity.

To diagnose symptomatic COPD, physicians utilize GOLD criteria based on the ratio of a patient's forced expiratory volume in one second (FEV₁) to their forced vital capacity (FVC) after bronchodilator use (8–10). After initial diagnosis (FEV₁ / FVC < 70%), disease progression is tracked via changes in percent predicted FEV₁ (%FEV₁), and is staged from GOLD 1 (> 80% predicted) to GOLD 4 (<30% predicted). Depending on the annual rate of FEV₁ decline, clinicians can then determine whether a smoker's lung function is declining rapidly ("decliner") or at a rate consistent with normal aging ("sustainer") (5,11). However, because spirometric measurements cannot differentiate between decliners and sustainers until *after* a significant functional decline has already occurred, they represent a much better tool for confirming disease progression than for predicting future functional decline (12).

Nevertheless, identifying smokers as decliners or sustainers is crucial for early treatment to prevent lung function decline (13). For example, Zhou and colleagues recently found that two years of treatment with tiotropium significantly improved FEV₁ compared to placebo in GOLD 1 and 2 COPD patients (14), as well as significantly decreasing the annual rate of FEV₁ decline in these same populations. The same effects were not seen in GOLD 3 and 4 COPD patients, however, indicating that early identification of functional decline is critical to stopping disease progression. The development of reliable prognostic indicators would allow for studies of treatment efficacy in patients at even earlier time points, before the appearance of symptoms, FEV₁ decline, and irreversible damage.

In order to better predict future FEV₁ deterioration associated with COPD development and progression, as well as to distinguish between decliners and sustainers earlier, we utilized wash-in multibreath imaging with inhaled hyperpolarized (HP) ³He. We measured regional

ventilation based on ^3He signal build-up in a series of images, and subsequently computed alveolar oxygen tension (P_{AO_2}), fractional ventilation (FV), and a localized measurement of airspace size and connectivity (ADC)—measurements sensitive to the severity of emphysema and small airway obstruction (15,16). By using this approach in a longitudinal human study, we sought to develop an imaging prediction model with which rapid decliners can be identified *before* a significant deterioration of FEV_1 occurs, something that non-imaging models cannot do. This powerful prediction model could allow rapid decliners to receive treatment sooner, thus effectively slowing or stopping lung function deterioration and COPD progression.

METHODS AND MATERIALS

Experimental procedures were as previously described in Hamedani et al. (17,18).

Subject Groups and Demographics

Subjects participating in this study were recruited through advertisements posted around the Hospital of the University of Pennsylvania from 2009 to 2013. A consecutive series of recruited subjects were subsequently placed into two groups: nonsmokers (0 pack years), and smokers with at least 20-pack years smoking history. For inclusion in either group, it was required that subjects were between the ages of 40 and 70, did not have any MRI contraindications, were not pregnant, and were not on supplemental oxygen therapy. 60 total subjects participated in the baseline visit for this study (15 nonsmokers and 45 smokers). Of the 45 smokers, the lowest reported pack year history was 20 and the highest was 50. Additionally, 12 of the smokers had a prior diagnosis of COPD. Participants in this study averaged 51 years of age. All subjects were consented for participation in the study and all experiments, as well as data collection, were approved by the local Institutional Review Board prior to beginning the study.

Of the 60 subjects who completed baseline imaging, 49 returned for an identical follow-up visit (12 non-smokers and 37 smokers); on average, this follow-up visit occurred 1.4 years after the baseline assessment (1 year at the least and 1.8 years at the most). If a patient experienced an FEV_1 decline of 100 mL/year between the two visits, they were deemed a “decliner.” Subjects who did not experience this rapid rate of FEV_1 decline were deemed “sustainers.” Because this study was concerned with predicting future lung function decline in current smokers, healthy nonsmokers were not included in the sustainer group, as rapid FEV_1 decline was not expected; inclusion of this group in the model would substantially increase its power to predict future decline, but would diminish its applicability to clinically relevant populations.

Clinical Tests

Prior to both baseline and follow-up imaging sessions, subjects performed pulmonary function tests (PFT) consisting of spirometry, plethysmography, and a measurement of diffusing capacity. All PFTs were performed by registered pulmonary function technologists at the Hospital of the University of Pennsylvania and were in accordance with ATS/ERS recommendations for acceptability and repeatability. The parameters of interest measured

were: FEV₁, FVC, %FEV₁, total lung capacity (TLC), residual volume (RV), and DL_{CO}. All PFTs were performed in accordance with ATS/ERS 2005 guidelines (19) at the Hospital of the University of Pennsylvania Pulmonary Function Laboratory. Additionally, all patients completed a six-minute walk test (6MWT) and a self-administered St. George Respiratory Questionnaire (SGRQ) prior to imaging. Although the questionnaire is divided into three sub-scores (symptoms, activity, and impact), only the overall score was used for comparison between subjects. For the 6MWT, the distance traveled, drop in oxygen saturation as determined by pulse oximetry, and dyspnea level were all monitored and recorded for inter-subject comparisons.

³He Imaging Markers

Following clinical testing at each visit, subjects participated in HP ³He MR imaging sessions. The three imaging markers measured during these imaging sessions were FV, P_AO₂, and ADC, which were used to assess lung functionality. These imaging markers are briefly described below.

- FV is defined as the ratio of the inspired gas entering the lungs upon inhalation to the functional residual capacity in the lungs following exhalation, known as the end expiratory volume (20). A regional equivalent to this global FV was determined through signal buildup of hyperpolarized gas over a multibreath wash-in protocol.
- P_AO₂ is defined as the partial pressure of oxygen in the lung parenchyma (17). Measured in Torr, it is calculated on a regional basis through decay of hyperpolarized gas signal during a breath-hold. Although P_AO₂ is generally expressed as a function of time due to the absorption of oxygen into the blood during a breath hold, measurements in this paper refer to the value assessed at end-inhalation.
- ADC is determined by measurements of the mean diffusion distance of ³He during approx. 1 ms of a breath hold, and is defined as the gas diffusion coefficient which would lead to an identical mean displacement in the absence of the restrictive alveolar walls and other structures. Measured ADC is directly related to alveolar size and connectivity, and is also sensitive to the size of peripheral airways (17,21).

Gas Delivery System and Imaging Session

HP ³He gas was delivered using a custom built, passively driven gas mixing and administration device, details of which were previously presented in Hamedani et al. (20). In essence, this device utilizes MRI-compatible differential pressure pneumotachometers, which independently report the flow rate of HP gas and oxygen during each inhalation. Once the target total volume is reached, pneumatic valves and the MRI scan are triggered. This subject-specific target volume was 12% of total lung capacity (TLC), as determined during PFT. Once the pneumatic valves were triggered to shut off the gas delivery, the subject then performed a voluntary breath hold. Subjects practiced the multibreath protocol with room air prior to each imaging session until they were able to perform it effectively. During imaging,

subjects performed seven inhalations (3 seconds duration) and exhalations (4 seconds duration) from functional residual capacity (FRC) to FRC+tidal volume, with a 1 second breath hold at the conclusion of the first six inhalations and a 12 second breath hold at the conclusion of the seventh inhalation, followed by exhalation and free breathing.

Approximately 1L of ^3He was hyperpolarized to ~30% during 15 hours of spin-exchange with optically pumped rubidium in a commercial polarizer (IGI 9600.He, GE Healthcare, Durham, North Carolina, USA). ^3He gas was then diluted with N_2 to yield the total gas volume required for seven breaths, transferred to the scanner in a TedlarTM bag, and further mixed with O_2 during inhalation to achieve normoxic conditions. The subject's blood oxygenation, heart rate, blood pressure, and respiratory rate were monitored by an on-site physician throughout each imaging session. Images of the gas delivery device and the signal build up over the seven inhalations / exhalations can be seen in Figure 1.

^3He MR Imaging

MR imaging was performed using a 1.5-T MRI (Avanto; Siemens Healthcare, Erlangen, Germany) and a custom-made, eight-channel chest coil (Stark Contrast, Erlangen, Germany) tuned to the 48.48-MHz ^3He resonance frequency placed over the subject's chest. Seven sets of back-to-back multislice end-inspiratory breath hold images were acquired during the wash-in and breath hold phase of breathing, and again during the wash-out phase when the patient was breathing room air.

During each round of image acquisition, end-inspiratory images acquired during wash-in and breath hold were used to calculate FV. Six coronal slices (thickness of ~20–25 mm) spanning the entire lung were acquired during the first, ~1 s breath hold (flip-angle: 5° , matrix size: 48×36 , TR/TE: 6.69/3.1 ms, 36 phase-encode lines). As shown in Figure 4B, FV was ultimately determined by assessing the regional signal build-up over the first six wash-in and wash-out cycles. Subjects were instructed to hold their breath for approximately 12 seconds after the seventh inhalation, during which simultaneous P_{AO_2} -ADC imaging was performed. Imaging parameters were similar to those used during FV imaging, except that an interleaved, diffusion-weighted sequence with TR/TE of 9.1/5.9ms was utilized. An expanded description of P_{AO_2} -ADC imaging has been previously presented by Hamedani et al. (17,18). Following the simultaneous P_{AO_2} -ADC imaging, subjects were instructed to exhale and breathe normally. Additional end-inspiratory images were acquired until a signal was no longer detected.

Gas Dynamic Models and Image Analysis

Following MR imaging, the lungs in each image were segmented (binned) into twelve isotropic regions of interest (ROIs) of approximately $\sim 500 \text{ cm}^3$ each through a semi-automated process operated by non-blinded, trained members of the investigative team (Figure 1). For every ROI, the mean and standard deviation (heterogeneity) of FV, P_{AO_2} , and ADC were calculated. The means and heterogeneity values are referred to hereafter as FV^{M} , FV^{H} , $\text{P}_{\text{AO}_2}^{\text{M}}$, $\text{P}_{\text{AO}_2}^{\text{H}}$, ADC^{M} and ADC^{H} . FV was calculated by acquiring signal build-up in a voxel-wise manner from the six identical back-to-back wash-in images acquired after HP ^3He inhalation. Signal intensities were fit using custom software

developed in MATLAB (MathWorks, Natick, Massachusetts, USA) to previously derived equations (17) in which the FV and an arbitrary overall signal scaling are allowed to vary. A third parameter, describing the local gas relaxation fraction per breath, is fixed at a physiologically plausible value and is iteratively refined as described below.

During the ~12 second breath hold following the seventh inhalation in the multibreath protocol, five additional images are acquired, including one with a diffusion-sensitizing bipolar gradient. The timing and acquisition parameters of these images are chosen such that the decay of signal can be separately attributed to RF-induced relaxation, O₂-induced relaxation, and ³He diffusion; a second fit using the modeled dependence on flip-angle α , P_AO₂, local O₂ uptake rate and ADC yields estimates of these parameters as well as a refined estimate of the fractional gas relaxation per breath. This updated value is then used to improve the accuracy of the FV fit, and iterative improvement continues until convergence is achieved. In practice, this process requires 3–4 iterations, at most (17).

Statistical Analysis

All statistical analyses were performed using R software. To compare the means of the non-imaging markers between cohorts, a one-way analysis of variance (ANOVA) was performed; in case of the imaging markers (with multiple spatial ROIs), a repeated measure ANOVA was used. We used a Type I error level of 0.05.

Two separate multivariate logistic regression analyses were performed in order to predict *decliner status* for both the non-imaging and imaging markers. In the case of repeated measures of the imaging markers (12 ROIs in each subject), we treated the ROIs as a random factor nested within each slice, which is specifically nested in the subject term to correct for the error term. To build our multivariate model, a stepwise methodology was used: first, a univariate logistic regression analysis was performed and covariates with $p_E < 0.25$ were picked for the initial multivariate model, in which the predictors remain if in the final model if $p_R < 0.25$. All the variables not used were added back one by one in the model and remain if $p_S < 0.3$ and/or significantly improve the model (either the coefficients or the Akaike Information Criterion). If confounding variables were identified during this step, these values remained in the model, as these would cause a significant change in parameter estimates (18). We tested the final prediction models with a leave-one-out cross validation (LOOCV). Finally, the sensitivity and specificity were calculated for prediction of decliner status, which was used to generate receiving operating character (ROC) curves.

RESULTS

Demographics and clinical (non-imaging) observations:

Demographic information for nonsmokers and smokers, as well as for smokers with COPD, is included in Table 1. Given their already compromised lung functionality, smokers with COPD were separated out; including them in the smoker group would have skewed the means and created large standard deviations. As a preliminary test of sensitivity to smoking-induced functional change, we analyzed potential differences in non-imaging markers between nonsmokers and smokers using a repeated measure ANOVA. These comparisons of

age, weight, height, SGRQ total score, 6MWT distance, RV/TLC, %FEV₁, FEV₁/FVC, and DLCO/VA are summarized in Figure 2A. Significant inter-group differences were found in SGRQ total score ($F(1,58)=9.92$, $p=0.003$), 6MWT distance ($F(1,53)=5.076$, $p=0.028$), %FEV₁ ($F(1,58)=9.12$, $p=0.004$), FEV₁/FVC ($F(1,58)=12.95$, $p<0.001$), and DLCO/VA ($F(1,56)=5.017$, $p=0.0291$); age, weight and height were not significantly different between groups ($p=0.273$, $p=0.388$ and $p=0.253$, respectively), while RV/TLC remained normal in the smokers ($F(1,58)=1.63$, $p=0.207$).

In order to determine whether or not lung function decline had occurred, the same clinical tests were performed on all subjects during follow up visits. Subjects who exhibited changes in FEV₁ > 100 mL/year (the approximate upper limit of normal, age-related decline (22)), were deemed decliners. 8 of the 37 smokers for which baseline and repeat visit measurements were available fell into this category (21.6%), and were well separated from the remaining 29 smokers in our dataset (Figure 3).

Having subcategorized smokers into sustainers and decliners using FEV₁, we then sought to use baseline, non-imaging measurements to predict subsequent FEV₁ decline. The box plots of Figure 2B provide a visual summary of these results, highlighting the minimal difference in all parameters between sustainers and decliners except SGRQ and DLCO, and suggesting that baseline FEV₁ assessments are a poor predictor of subsequent FEV₁ decline. It should be noted that in order to expand the range of baseline FEV₁ measurements examined, 12 subjects who had been previously diagnosed with COPD were included in this study. For this group of subjects, COPD GOLD classification was a poor predictor of the rate of future decline—confirming what previous studies have shown (23). As with the comparison between smokers and nonsmokers, the heterogeneity of individual measurements in both sustainers and decliners is such that no single parameter can be used to determine decliner status.

To find the best predictors of lung function decline, we performed a step-wise logistic regression analysis on all baseline non-imaging makers. An initial univariate logistic regression analysis (Table 1A) indicated that DLCO/VA, SGRQ total score, and FEV₁/FVC enter the initial multivariate logistic regression analysis ($p_E<0.25$). Of these parameters, the subjective SGRQ appeared to be the best marker of FEV₁ decline, suggesting subjects' awareness of their own compromised functionality. To more rigorously quantify the diagnostic power of these non-imaging markers—with the ultimate goal of developing a non-imaging model to predict future FEV₁ decline—putative predictors (Age, Height, Pack Years, DLCO/VA, and FEV₁/FVC) were then added back to the model. Stepwise testing, removal and re-introduction of variables based on p_E and p_R (18) resulted in a three-factor model depending on SGRQ total score, DLCO/VA and 6MWT distance, the latter of which re-entered the model as a confounding variable after initially being dropped. When re-introduced, 6MWT distance significantly changed the model by reducing the p-values of both SGRQ total score ($p=0.039$) and DLCO/VA ($p=0.004$). Final estimates for decliner status, together with the statistical results of the multivariate logistic regression analysis, are shown in Table 2B. To validate the regression models, a leave-one-out cross validation analysis was performed for all subjects in the sustainer cohort vs. the decliner cohort. The

sensitivity and specificity of this non-imaging prediction model are presented as an ROC curve in Figure 3: at the operating point, they were 75.0% and 68.3%, respectively.

Imaging Studies:

Figure 5 shows representative FV, P_{AO_2} , and ADC maps from a decliner subject acquired during HP ^3He MR imaging, with the anterior slices of the lung on the far left, progressing down to the posterior slices on the far right. Each of these imaging markers was quantified over the 12 ROIs for each slice from each subject, which generated means (M) and standard deviations (or heterogeneity (H)) of FV, P_{AO_2} , and ADC specific to each of the 12 ROIs in every slice, and allowing us to assess regional lung functionality and its uniformity in every subject. The box plots in Figure 6A summarize the differences observed in each marker between smokers and nonsmokers.

As with the non-imaging markers, a repeated measure ANOVA was then performed on the regional (12 ROIs) imaging parameters (mean and heterogeneity of FV, P_{AO_2} and ADC) in order to highlight markers sensitive to smoking-induced changes to the lung: heterogeneity of P_{AO_2} ($P_{AO_2}^H$, [F(1,59)=17.23, $p < 0.001$]) showed the greatest change in smokers compared to nonsmokers, indicating that it could potentially serve to differentiate between these two groups. Additionally, we see that the mean fractional ventilation (FV^M) is lower in smokers ([F(1,59)=15.65, $p < 0.001$]). No other metric showed a significant difference between smoking and nonsmoking cohorts; although ADC is increased in a portion of smokers, both mean and standard deviation are the same between the two cohorts (F(1,59)=2.537, $p=0.117$ and F(1,59)=3.218, $p=0.185$, respectively). Additionally, $P_{AO_2}^M$ (F(1,59)=0.074, $p=0.786$) and FV^H (F(1,59)=0.125, $p=0.725$) were not significantly different between smokers and nonsmokers.

Again mirroring our approach to the non-imaging markers studied, we next sought to use baseline imaging-based metrics to predict future functional decline. Using the same definition of sustainers and decliners (> 100 mL/year decline in FEV1), we followed an identical methodology to build a multivariate logistic regression model for predicting the decliner status of an individual lung region. To account for the repeated measures in a single subject, we used a mixed model containing the nested ROIs/Slice/Subjects as a random term. Imaging markers with an initial univariate $p_E < 0.25$ were entered into the model, as shown in Table 3A (all but FV^H). Reduced p-values then required all covariates except ADC^H and $P_{AO_2}^H$ to be dropped. FV^H was subsequently added back to the model, as it led to both a significantly decreased AIC and an improvement in coefficients. A summary of the final multivariate prediction model is given in Table 3B.

These results were tested in a LOOCV to generate the ROC curves of Figure 7, which show the sensitivity and specificity of individual (Figure 7A-C) and combined (Figure 7D) imaging parameters for detecting an alteration in ROIs belonging to a representative decliner subject. Individual ROC curves of ADC^H , $P_{AO_2}^H$, and FV^H are also shown in order to demonstrate the substantially increased ability to distinguish the abnormalities in each ROI offered by a multifaceted regional prediction model containing different functional and structural parameters of lung health instead of single parameters on their own. The combined ROC curve displays a sensitivity of 82.9% and a specificity of 80.6% at its operating point,

indicating that this combined prediction model of ADC^H , $P_{AO_2}^H$, and FV^H can differentiate subclinical changes in the ROIs from a decliner's lung quite accurately. It is important to note, however, that these ROC curves are not directly comparable to the ROC curve seen in Figure 4, as they only address the decliner status of an individual lung region. To facilitate a meaningful comparison, we therefore characterized each subject in terms of the number of true positive regions, defined as regions predicted to come from a decliner, in order to find an optimal cutoff for separating decliners from sustainers. The ROC curves generated using LOOCV following a univariate logistic regression model (Coeff.=0.754, SE=0.308, $z=2.447$, $p=0.014$) are shown in Figure 8.

The overall value of our HP ^3He MR imaging prediction model can be appreciated by comparing its sensitivity and specificity to those of the non-imaging prediction model of decliner status discussed above. At its operating point, the imaging model increased sensitivity and specificity from 75% to 88% and from 68% to 83%, respectively—resulting in an overall increase of 25% in the area under the curve (AUC).

DISCUSSION

In this study, we developed an improved method for predicting rapid lung function decline in smokers before it occurs (24), and which has the potential to be beneficial for the diagnosis and treatment of COPD. Unlike the non-imaging markers, the imaging markers studied here enabled us to effectively detect future decliners with high specificity and sensitivity. If validated in a larger study with a training dataset, this method could serve as a valuable clinical tool for determining which patients are at the highest risk for developing COPD, allowing clinicians to either begin treatment or strongly recommend that they stop smoking. Additionally, early-stage COPD patients could potentially begin treatment with tiotropium before the disease has progressed to the point at which point this medication can no longer improve and/or preserve lung functionality (14).

A closer analysis of our non-imaging parameters shows that the smoking group displayed a slightly altered mean and a greatly increased standard deviation across all measurements compared to non-smokers (Figure 1). Although measurement distributions differ significantly, this indicates that individual measurements on their own may be non-specific in predicting smoking status—a finding consistent with the well-known heterogeneity of response to cigarette smoke (18). Surprisingly, neither %FEV1 or FEV1/FVC entered this model at all, suggesting that although these measurements are indicative of COPD status, they may not help predict future decline. It is worth noting that this model's relatively low sensitivity shows that the combination of non-imaging metrics will often present with seemingly normal results even for subjects who will experience significant lung function decline in the near future. Though higher, the model's specificity must be interpreted in light of the low prevalence of decliner status, which means that its positive predictive value for future disease progression is also poor.

Turning to our imaging measurements, it is important to note that mean fractional ventilation is expected to reflect the relationship between administered gas volume and TLC. It is therefore somewhat surprising that FV^M appears systematically different in the smoking

population imaged here. This may indicate either a shift in the FV^M/TLC relationship due to regions of air trapping (25,26), or an FV measurement bias caused by gas transport within the lung not accounted for in the FV model (20). Additionally, in comparing the individual ROC curves of Figure 7A-C, we see that, of the three imaging parameters included in the final model, ADC makes the smallest contribution to its predictive power. Since ADC is most directly sensitive to emphysematous remodeling (21,27), this suggests that the current extent of emphysema is a poorer predictor of disease progression than had been previously postulated (28–30). Instead, our model's more robust dependence on measures of ventilation (FV^H) and oxygenation ($P_AO_2^H$) suggests that small airway obstruction was more closely associated with subsequent functional decline in subjects imaged in this study. This finding is consistent with results previously presented in a large scale study on COPD patients utilizing CT imaging (31). Although PFTs can also be sensitive to small airway obstruction, they lack the regional sensitivity of our HP-based measurements. Because our imaging model predicts decliner status based on the number of regions (or fraction of the lung) displaying abnormal function rather than on whole-lung mean or standard deviation, its greater predictive value relative to PFTs suggests that a lung which is uniformly but modestly compromised is more prone to decline than an organ with locally severe disease which is otherwise normal.

The primary limitation of this study arises from the fact that only 8 of the 60 subjects imaged experienced the > 100 ml/year FEV1 decline necessary for characterization as decliners, making it impossible to support a separate training dataset. Ideally, the pilot study presented here would be used as such, so that its hypotheses could be further tested using separate, similarly-sized cohorts. Additionally, of the 15 nonsmokers and 45 smokers who underwent baseline imaging, only 12 nonsmokers and 37 smokers returned for their follow-up visit. While this study only compared baseline visit statistics, more decliners could have potentially been identified had all subjects undergone a second round of imaging. It is also important to note that the power of our logistic regression model is inevitably reduced by the fact that only 8 subjects experienced an FEV1 decline of 100 mL/year. Given the practical limitations of recruiting enough “decliner” subjects and our inability to predict this decline prior to the conclusion of this work, we performed the LOOCV along with the logistic regression. Although this mitigated the problem posed by the low number of decliners, one well-known drawback to performing a LOOCV in combination with a logistic regression is the tendency for this methodology to exaggerate the AUC in the ROC curves. Additionally, there was no attempt to determine whether a better set of predictors could yield a more effective prediction model while performing the LOOCV. As a result, these prediction models contain an expected bias. However, because this bias is a result of the statistical analyses performed, it does not undermine the findings presented, as all statistical analyses on both non-imaging measurements and imaging markers were performed in the same manner.

In this work, we developed a prediction model that utilizes imaging markers generated from HP ^3He MR imaging to predict future FEV1 decline in smokers with high specificity. This model proved to be more effective than one based on non-imaging measurements, which included those performed during standard clinical pulmonary assessments. Once validated by further studies, the imaging model described here could be used to determine whether a

smoker will experience rapid FEV1 decline/develop COPD, providing a valuable clinical tool for early treatment and intervention. Although more speculatively, this work also suggests that decliners exhibited signs of small airway obstruction (as determined by inclusion of FV^H and $P_{AO_2}^H$ in the model) more often than they do emphysematous remodeling (as determined by the exclusion of ADC^M from the model), potentially indicating that small airway obstruction actually plays a larger role in FEV1 decline than emphysema. Finally, our results also suggest that modest but widespread structural or functional change in the lung is more predictive of subsequent decline than localized, severe alterations. Future research should seek to further elucidate the link between small airway obstruction and FEV1 decline using HP noble gas MR imaging.

Acknowledgements

This work was funded by NIH Grant RO1-HL127969.

References

1. Currie GP. ABC of COPD, 3rd Edition. 3rd ed Wiley-Blackwell; 2017 <https://www.wiley.com/en-us/ABC+of+COPD%2C+3rd+Edition-p-9781119212850>. Accessed February 28, 2018.
2. WHO. WHO | Chronic obstructive pulmonary disease (COPD). Chronic Obstr. Pulm. Dis. COPD <http://www.who.int/respiratory/copd/en/>. Accessed February 28, 2018.
3. Fletcher C, Peto R. The natural history of chronic airflow obstruction. *Br Med J*. 1977;1(6077):1645–1648. [PubMed: 871704]
4. Fletcher C, Peto R, Tinker C, Speizer F. The natural history of chronic bronchitis and emphysema. Oxford: Oxford University Press 1976.
5. Vestbo J, Edwards LD, Scanlon PD, et al. Changes in forced expiratory volume in 1 second over time in COPD. *N Engl J Med*. 2011;365(13):1184–1192. [PubMed: 21991892]
6. Ruppert K Biomedical imaging with hyperpolarized noble gases. *Rep Prog Phys Phys Soc G B*. 2014;77(11):116701.
7. Albert MS, Balamore D. Development of hyperpolarized noble gas MRI. *Nucl Instrum Methods Phys Res Sect Accel Spectrometers Detect Assoc Equip*. 1998;402:441–453.
8. Rodriguez-Roisin R, Rabe KF, Vestbo J, Vogelmeier C, Agustí A, all previous and current members of the Science Committee and the Board of Directors of GOLD (goldcopd.org/committees/). Global Initiative for Chronic Obstructive Lung Disease (GOLD) 20th Anniversary: a brief history of time. *Eur Respir J*. 2017;50(1).
9. Woodruff PG, Barr RG, Bleecker E, et al. Clinical Significance of Symptoms in Smokers with Preserved Pulmonary Function. *N Engl J Med*. 2016;374(19):1811–1821. [PubMed: 27168432]
10. Vogelmeier CF, Criner GJ, Martinez FJ, et al. Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Lung Disease 2017 Report. GOLD Executive Summary. *Am J Respir Crit Care Med*. 2017;195(5):557–582. [PubMed: 28128970]
11. Nishimura M, Makita H, Nagai K, et al. Annual change in pulmonary function and clinical phenotype in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2012;185(1):44–52. [PubMed: 22016444]
12. Wise RA. The value of forced expiratory volume in 1 second decline in the assessment of chronic obstructive pulmonary disease progression. *Am J Med*. 2006;119(10 Suppl 1):4–11.
13. Decramer M, Cooper CB. Treatment of COPD: the sooner the better? *Thorax*. 2010;65(9):837–841. [PubMed: 20805184]
14. Zhou Y, Zhong N-S, Li X, et al. Tiotropium in Early-Stage Chronic Obstructive Pulmonary Disease. *N Engl J Med*. 2017;377(10):923–935. [PubMed: 28877027]

15. Hamedani H, Kadlecsek SJ, Emami K, et al. A multislice single breath-hold scheme for imaging alveolar oxygen tension in humans. *Magn Reson Med*. 2012;67(5):1332–1345. [PubMed: 22190347]
16. Yu J, Law M, Kadlecsek S, et al. Simultaneous measurement of pulmonary partial pressure of oxygen and apparent diffusion coefficient by hyperpolarized ³He MRI. *Magn Reson Med*. 2009;61(5):1015–1021. [PubMed: 19230017]
17. Hamedani H, Kadlecsek S, Xin Y, et al. A hybrid multibreath wash-in wash-out lung function quantification scheme in human subjects using hyperpolarized ³He MRI for simultaneous assessment of specific ventilation, alveolar oxygen tension, oxygen uptake, and air trapping. *Magn Reson Med*. 2017;78(2):611–624. [PubMed: 27734519]
18. Hamedani H, Kadlecsek SJ, Ishii M, et al. Alterations of regional alveolar oxygen tension in asymptomatic current smokers: assessment with hyperpolarized (³)He MR imaging. *Radiology*. 2015;274(2):585–596. [PubMed: 25322340]
19. American Thoracic Society European Respiratory Society. ATS/ERS recommendations for standardized procedures for the online and offline measurement of exhaled lower respiratory nitric oxide and nasal nitric oxide, 2005. *Am J Respir Crit Care Med*. 2005;171(8):912–930. [PubMed: 15817806]
20. Hamedani H, Clapp JT, Kadlecsek SJ, et al. Regional Fractional Ventilation by Using Multibreath Wash-in (³)He MR Imaging. *Radiology*. 2016;279(3):917–924. [PubMed: 26785042]
21. Diaz S, Casselbrant I, Piitulainen E, et al. Hyperpolarized ³He apparent diffusion coefficient MRI of the lung: reproducibility and volume dependency in healthy volunteers and patients with emphysema. *J Magn Reson Imaging JMRI*. 2008;27(4):763–770. [PubMed: 18344208]
22. Kretschman D, Gao W, Dupuis J, Latourelle J, O'Connor G. Rate Of FEV1 Decline In Healthy Adults: Defining The Upper Limit Of Normal In The Framingham Heart Study A66 Model Mech GAS Exch. American Thoracic Society; 2012 p. A2056–A2056 10.1164/ajrccm-conference.2012.185.1_MeetingAbstracts.A2056. Accessed March 30, 2018.
23. Kim J, Yoon HI, Oh Y-M, et al. Lung function decline rates according to GOLD group in patients with chronic obstructive pulmonary disease. *Int J Chron Obstruct Pulmon Dis*. 2015;10:1819–1827. [PubMed: 26379432]
24. Krishnan JK, Martinez FJ. Lung function trajectories and chronic obstructive pulmonary disease: current understanding and knowledge gaps. *Curr Opin Pulm Med*. 2018;24(2):124–129. [PubMed: 29206658]
25. Altes TA, Powers PL, Knight-Scott J, et al. Hyperpolarized ³He MR lung ventilation imaging in asthmatics: preliminary findings. *J Magn Reson Imaging JMRI*. 2001;13(3):378–384. [PubMed: 11241810]
26. McMahon CJ, Dodd JD, Hill C, et al. Hyperpolarized ³helium magnetic resonance ventilation imaging of the lung in cystic fibrosis: comparison with high resolution CT and spirometry. *Eur Radiol*. 2006;16(11):2483–2490. [PubMed: 16871384]
27. Salerno M, de Lange EE, Altes TA, Truwit JD, Brookeman JR, Mugler JP. Emphysema: hyperpolarized helium 3 diffusion MR imaging of the lungs compared with spirometric indexes--initial experience. *Radiology*. 2002;222(1):252–260. [PubMed: 11756734]
28. Mohamed Hoessein FAA, de Hoop B, Zanen P, et al. CT-quantified emphysema in male heavy smokers: association with lung function decline. *Thorax*. 2011;66(9):782–787. [PubMed: 21474499]
29. Mohamed Hoessein FAA, van Rikxoort E, van Ginneken B, et al. Computed tomography-quantified emphysema distribution is associated with lung function decline. *Eur Respir J*. 2012;40(4):844–850. [PubMed: 22323577]
30. Tanabe N, Muro S, Tanaka S, et al. Emphysema distribution and annual changes in pulmonary function in male patients with chronic obstructive pulmonary disease. *Respir Res*. 2012;13:31. [PubMed: 22512922]
31. Bhatt SP, Soler X, Wang X, et al. Association between Functional Small Airway Disease and FEV1 Decline in Chronic Obstructive Pulmonary Disease. *Am J Respir Crit Care Med*. 2016;194(2):178–184. [PubMed: 26808615]

Figure 1:

Study design and data analysis flow chart depicting the steps from data collection to statistical analysis in this study. Data collection occurred during the MRI session, after which images were registered (adjusted so that all were uniform in size and alignment). Once the images were adjusted, the analysis of ADC, $P_{\text{A}}\text{O}_2$, and FV began with the binning of images into the 12 ROI's (binned voxels) of 500 cm^3 each. Statistical analyses could then be successfully run on this data.

Figure 2:

Box plot representations of non-imaging markers comparing (A) smokers vs. non-smokers and (B) sustainers vs. decliners. The line within each box represents the median, while the circle in the middle represents the mean. Values for FEV₁/FVC, %FEV₁, DLCO/VA, and RV/TLC were determined via PFT. SGRQ overall represents the total overall scores from the St. George Respiratory Questionnaires. RV= residual volume, TLC= total lung capacity, 6MWT= 6-minute walk test.

Figure 3:

FEV₁ drop between baseline and follow up imaging session. The bolded black lines indicate a substantial FEV₁ decline of at least 100 mL/yr. In total, 8 subjects experienced significant FEV₁ decline between baseline and follow-up visits.

Figure 4:

ROC curve generated for the results of the multivariate logistic regression, which entered the non-imaging prediction model for future FEV1 decline between sustainers and decliners. The three non-imaging markers that entered this model were SGRQ total score, DLCO/VA, and 6MWT. At the operating point in this model sensitivity=75.0% and specificity= 68.3%.

Figure 5:
Complete regional lung function assessment consisting of ventilation (FV), alveolar oxygen tension (P_{AO_2}), and ADC in a representative subject. Images were acquired from one asymptomatic smoker ($FEV_1/FVC=72$; $DL_{CO}=30.79$, $\%FEV_1=93$).

Figure 6:

Boxplots of the imaging markers comparing (A) smokers vs. nonsmokers and (B) sustainers vs. decliners. The line within each box represents the median, while the circle in the middle represents the mean. Each point on the boxplots represents the specified value calculated from one of the 12 ROI's from each slice of MR lung images. All imaging markers were generated during HP ^3He MR imaging sessions. H= heterogeneity (standard deviation); M= mean.

Figure 7:

ROC curves of the regional measures (A) ADC^H (Sensitivity=82.4, Specificity=39.6), (B) $P_{AO_2}^H$ (Sensitivity=64.0, Specificity=96.9), (C) FV^H (Sensitivity=60.7, Specificity=90.8).

These three parameters entered into the imaging prediction model of future FEV1 decline used to differentiate sustainers from decliners. (D) ROC curve of a multifaceted measure of ADC^H , $P_{AO_2}^H$, and FV^H in the imaging prediction model (Sensitivity=82.9, Specificity=80.8).

Figure 8:

ROC curve of the predictive imaging model, with the optimal cut-off for differentiating sustainers from decliners (solid block line), and the non-imaging model (red line). At its operating point, the imaging prediction model has a sensitivity of 87.5% and a specificity of 82.8%. This model also possesses an AUC of 89.7%.

Table 1:

Demographic information for nonsmokers, smokers, and subjects in the smokers group with COPD. Values are presented as mean $\pm$ standard deviations. RV= residual volume, TLC= Total lung capacity.

	Nonsmokers (n=15)	Smokers (n=33)	Smokers with COPD (n=12)
Demographics			
Age	52.9 $\pm$ 5.9	49.9 $\pm$ 6.9	52.4 $\pm$ 8.2
Height (in)	67.8 $\pm$ 3.4	69.0 $\pm$ 3.1	68.6 $\pm$ 3.5
Weight (lb)	174.6 $\pm$ 29.5	183.9 $\pm$ 34.9	181.65 $\pm$ 36.9
Body mass index (kg/m ²)	26.7 $\pm$ 4.2	27.2 $\pm$ 4.04	27.0 $\pm$ 5.3
Smoking (Pack Years)	0 $\pm$ 0	29.1 $\pm$ 10.3	38.1 $\pm$ 16.4
PFTs			
FVC(L)	4.09 $\pm$ 0.96	4.17 $\pm$ 1.1	3.55 $\pm$ 1.0
FEV1 (L)	3.35 $\pm$ 0.78	3.20 $\pm$ 0.72	2.03 $\pm$ 0.62
FEV1/FVC (%)	81.7 $\pm$ 2.7	77.5 $\pm$ 4.5	57.3 $\pm$ 5.3
RV/TLC (%)	30.6 $\pm$ 4.5	31.1 $\pm$ 7.5	40.5 $\pm$ 8.1
DLCO (mL/min/mm Hg)	27.4 $\pm$ 5.1	25.3 $\pm$ 6.3	18.4 $\pm$ 5.3
Percentage predicted FEV1	106.1 $\pm$ 11.1	98.1 $\pm$ 15.1	65.7 $\pm$ 12.5
Percentage predicted FVC	101.7 $\pm$ 9.2	101.4 $\pm$ 18.8	90.2 $\pm$ 14.5
Percentage Predicted DLCO	107.4 $\pm$ 12.5	98.2 $\pm$ 22.9	81.1 $\pm$ 21.9
Clinical Tests			
Distance walked in 6MWT (m)	551.8 $\pm$ 120.8	485.1 $\pm$ 70.0	506.3 $\pm$ 72.6
SGRQ Overall Score	1.91 $\pm$ 2.7	11.8 $\pm$ 12.7	27.6 $\pm$ 22.10

Table 2:

(A) All non-imaging markers were assessed during clinical tests prior to MR imaging. %FEV1, DLCO/VA, RV/TLC, and FEV1/FVC were determined via PFT. (B) Non-imaging markers selected for the multivariate logistic regression were those values that had a p-value <0.25 from the univariate logistic regression.

A) Univariate Logistic Regression of Non-imaging Markers Sustainers vs. Decliners					
Variable	Coeff.	SE	z	p	AIC
Age	0.034	0.043	-1.299	0.194	68.6
Height	0.197	0.112	1.764	0.078	65.6
Weight	-0.002	0.009	-0.234	0.815	69.1
Pack Years	0.064	0.023	2.793	0.005	58.7
6MWT	-0.001	0.001	-1.148	0.251	60.3
SGRQ	0.049	0.020	2.397	0.017	62.2
%FEV1	-0.012	0.015	-0.779	0.436	68.6
DLCO/VA	-2.011	0.590	-3.410	0.001	45.9
RV/TLC	0.051	0.039	1.304	0.192	67.5
FEV1/FVC	-0.102	0.033	-3.070	0.002	58.2
B) Multivariate Logistic Regression of Non-imaging Markers Sustainers vs. Decliners					
Variable	Coeff.	SE	z	p	AIC
SGRQ	0.076	0.037	2.055	0.039	
DLCO/VA	-2.039	0.705	-2.893	0.004	40.8
6MWT	0.002	0.002	1.118	0.263	

Table 3:

(A) All imaging markers were generated during HP ^3He MR imaging. (B) The values that entered the multivariate logistic regression were those values that had a p-value <0.25 from the univariate logistic regression.

A) Univariate Logistic Regression of Non-imaging Markers Sustainers vs. Decliners					
Variable	Coeff.	SE	z	p	AIC
ADC^M	0.832	0.865	1.181	0.236	75.0
ADC^H	-1.964	0.724	-2.711	0.007	63.6
P_aO₂^m	-0.915	0.554	-1.652	0.099	73.3
P_aO₂^H	-2.903	1.237	-2.348	0.019	61.2
FV^M	-1.576	0.757	-2.083	0.037	71.4
FV^H	0.131	0.411	0.320	0.749	76.6
B) Multivariate Logistic Regression of Imaging Markers Sustainers vs Decliners					
Variable	Coeff.	SE	z	p	AIC
ADC^H	-3.668	2.290	-1.602	0.109	
P_aO₂^H	-3.113	1.864	-1.671	0.095	55.0
FV^H	1.446	1.348	1.072	0.284	