Correcting Nonpathological Variation in Longitudinal Parametric Response Maps of CT Scans in COPD Subjects: SPIROMICS

Antonio Fernández-Baldera[1], **Charles R. Hatt**[2], **Susan Murray**[3], **Eric A. Hoffman**[4], **Ella A. Kazerooni**[1], **Fernando J. Martinez**[5], **MeiLan K. Han**[6], and **Craig J. Galbán**[1]

[1]Department of Radiology, University of Michigan, Ann Arbor, MI

[2]Imbio, LLC. Minneapolis, MN

[3]Department of Public Health, University of Michigan, Ann Arbor, MI

[4]Departments of Radiology and Biomedical Engineering, University of Iowa, IA

[5]Department of Medicine, Cornell University, NY

[6]Department of Internal Medicine, University of Michigan, Ann Arbor, MI

Abstract

Small airways disease (SAD) is one of the leading causes of airflow limitations in patients diagnosed with chronic obstructive pulmonary disease (COPD). Parametric response mapping (PRM) of computed tomography (CT) scans allows for the quantification of this previously invisible COPD component. Although PRM is being investigated as a diagnostic tool for COPD, variability in the longitudinal measurements of SAD by PRM has been reported. Here, we show a method for correcting longitudinal PRM data because of non-pathological variations in serial CT scans. In this study, serial whole-lung high-resolution CT scans over a 30-day interval were obtained from 90 subjects with and without COPD accrued as part of SPIROMICS. It was assumed in all subjects that the COPD did not progress between examinations. CT scans were acquired at inspiration and expiration, spatially aligned to a single geometric frame, and analyzed using PRM. By modeling variability in longitudinal CT scans, our method could identify, at the voxel-level, shifts in PRM classification over the 30-day interval. In the absence of any correction, PRM generated serial percent volumes of functional SAD with differences as high as 15%. Applying the correction strategy significantly mitigated this effect with differences ~1%. At the voxel-level, significant differences were found between baseline PRM classifications and the follow-up map computed with and without correction ($P < .01$ over GOLD). This strategy of accounting for nonpathological sources of variability in longitudinal PRM may improve the quantification of COPD phenotypes transitioning with disease progression.

Corresponding Author: Craig J. Galbán, PhD, Department of Radiology, University of Michigan, BSRB, Room A506, 109 Zina Pitcher Place, Ann Arbor, MI 48109-2200; cgalban@med.umich.edu.

Keywords

COPD; parametric response mapping; longitudinal; computed tomography

INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is a leading cause of morbidity, mortality, and burden on the world's health and financial systems (1, 2). Advances in the clinical management of patients with COPD have led to an improved understanding of the multitude COPD phenotypes. It has been postulated that a spectrum of pathological processes may result in unique progression patterns among these patients. Extensive research has been devoted toward identifying surrogate biomarkers of disease progression with a strong emphasis on noninvasive imaging techniques and analytical approaches (3).

Parametric response mapping (PRM) is an analytical approach that, when applied to spatially aligned high-resolution computed tomography (HRCT) scans, allows both visualization and quantification of lung parenchyma affected by small airways disease (SAD), even when only emphysema is visibly observed (4). This technique quantifies a previously occult component of COPD and can be applied to retrospective HRCT data. Included in various NIH-funded clinical trials on COPD (5, 6), PRM of functional SAD (fSAD) has been demonstrated as an independent indicator of clinically relevant outcome measures (7). More recent studies have identified PRM as a surrogate of spirometric decline in COPD (7) and also a means for identifying and monitoring the onset of bronchiolitis obliterans syndrome in bone marrow and lung transplant recipients (8–10). In a preliminary study, PRM was evaluated as a marker for monitoring change in disease classification (ie, normal, fSAD, and emphysema) from subjects accrued as part of SPIROMICS (5). In this study, "voxel-based tracking," a method for evaluating longitudinal changes in PRM classification at the voxel level, has been used. Although this approach when applied to PRM shows promise at providing local disease progression, variability in Hounsfield unit (HU) values from uncontrollable sources (eg, scanner noise, patient breathing level, and image registration) may result in shifts in voxel PRM classification that are not related to alterations in disease state (6).

Various studies have demonstrated the efficacy of PRM as a diagnostic and prognostic indicator of decline in pulmonary function and COPD severity. Nevertheless, the use of PRM to monitor COPD progression has shown a high sensitivity of voxel classification to HU variability between longitudinal CT examinations resulting in erroneous results. The purpose of this study was to present a strategy to mitigate the effects of nonpathological HU variability on voxel classification for analyzing COPD progression using PRM.

METHODOLOGY

Study Population

All clinical procedures were conducted under an institutional review board-approved protocol, and all subjects involved provided written informed consent. In total, 90 subjects

(age range at baseline, 40–80 years), with paired volumetric inspiratory and expiratory HRCT scans and clinical examinations at a 30-day interval, were prospectively accrued as part of the Repeatability and Replicate Substudy of SPIROMICS (11). Subjects evaluated included smokers with a smoking history of 20 pack-years and GOLD (Global Initiative for Chronic Obstructive Lung Disease) scores across the scale including 0, 1, 2, 3 and 4 (11–13) and never-smokers (79 smokers and 11 never-smokers, respectively). Postbronchodilator forced expiratory volume at 1 second (FEV1) was determined by spirometry at each time point. In our set of subjects, exclusion criteria included intolerance to bronchodilators, body mass index (BMI) >40 kg/m^2 at baseline, presence of non-COPD obstructive lung disease, diagnosis of unstable cardiovascular disease, lung surgery, or presence of metal in the chest that might affect chest CT interpretation. Further, 13 subjects from this cohort have been previously used to define thresholds that indicate disease-provoked changes in PRM metrics (5).

Computed Tomography

Whole-lung volumetric multidetector HRCT scans were acquired for all 90 subjects using the SPIROMICS imaging protocol (13). The current of 120 kVP was adjusted to meet the CT dose index volume targets for inspiration and expiration by making use of 3 settings— large (BMI > 30 kg/m^2), medium (BMI, 20–30 kg/m^2), and small (BMI < 20 kg/m^2)—with vendor-specific reconstruction kernels (Standard, B, B35, FC03) (11). In this study, HRCT data reconstructed using the "GE standard" kernel were analyzed. Quantitative HRCT data were presented in HU values, in which stability of CT measurements for each scanner was monitored on a monthly basis by use of the COPDGene phantom (14). For reference, ambient air and water attenuation values should be −1000 and 0 HU, respectively. Of the 90 subjects, rescanning using a different scanner was conducted among 11 subjects and that using a different field of view (FOV> 5%) from their original scan was conducted among 22 subjects. To reduce scanner noise, a 3^3 median filter was applied to all CT scans before processing and analysis.

Parametric Response Mapping

Lung segmentation and image registration to a single geometric frame (ie, baseline expiration CT scan) were performed on all paired CT data using Lung Density Analysis (LDA) software (Imbio, LLC, Minneapolis, MN). Classification of individual voxels was performed using in-house algorithms developed using MATLAB version 2015b (MathWorks, Inc, Natick, MA). Details on the PRM analysis have been previously reported (4). The nomenclature of these measures for normal lung parenchyma, fSAD, and emphysema includes PRM^{Normal}, PRM^{fSAD}, and PRM^{Emph}, respectively. Additional details are provided in the online supplemental Methods.

Correction Strategy for Longitudinal PRM

After establishing that the difference in HU between interval examinations has a quasi-normal distribution (online supplemental Methods), we determined the variance using the serial inspiration and expiration voxel data. This approach is analogous to previous works on voxel-to-voxel therapeutic response assessment in cancer (15, 16). These data were plotted on a Cartesian coordinate system with the *x* and *y* axes denoted as the baseline and follow-

up, respectively. Using principal component analysis, the data were transformed to the axes of primary and secondary variance (principal and secondary eigenvectors, respectively). Next, a linear fit along the principal eigenvector was performed with the subsequent residuals mapped into the second eigenvector axis. Residuals were used to calculate the 95% confidence interval of the fit (95% CI). The value of the confidence interval was transformed back to the original image space and defined as the index of measurement variability (IMV). This procedure was performed among all subjects and CT breath-hold examinations at both inspiration and expiration. To account for HU dependence on IMV, a cumulative exponential model was applied to all voxels in the CT data:

$$IMV(x) = V[1 - e^{r(x+1000)}] + \delta,$$

where $V = A - \delta$, A is the maximum amplitude of IMV, δ is the minimum amplitude of IMV (± 30 HU), r is a rate constant, and x is the voxel HU value. Model derivation and calculation of parameter values are provided in the online supplemental Methods. This functional form of IMV was incorporated into the correction strategy to account for variations in voxel variance.

Figure 1 shows an illustration of the correction strategy. In step 1, we calculated the difference and average maps, ie, ΔHU and <HU>, respectively, between serial examinations. In step 2, we applied the following logical statement to each voxel of the baseline scan: if |ΔHU| < IMV(<HU>), then voxel =1, or else voxel = 0. This binary map was multiplied to the baseline scan, whereas the inverse was multiplied to the follow-up scan. In step 3, a composite follow-up CT scan was generated by summing the masked baseline and follow-up scans from step 2. Finally, in step 4, we calculated PRM (4). All data processing was performed using MATLAB version 2015b (MathWorks, Inc).

Statistical Analysis

Differences in subject age, height, weight, BMI, FEV1 (percent predicted), lung volumes, and percent lung volumes of PRM metrics at both interval examinations were assessed using the 1-way ANOVA controlled for multiple comparisons (Bonferroni post hoc test). Association between gender population and GOLD was assessed using the log likelihood ratio test. Temporal changes in lung volumes, relative volume of PRM metrics, and FEV1 were assessed using the Wilcoxon signed rank test. Differences in the percent agreement, ie, sum of voxels with like-PRM class normalized to total lung voxels, between uncorrected and corrected PRM at follow-up were also analyzed using the Wilcoxon signed rank test. Differences in percent agreement were also analyzed over GOLD stratums using the Kruskal–Wallis test. Percent agreement was computed using MATLAB version 2015b (MathWorks, Inc.). Statistical analyses were conducted with SPSS version 2.1 (IBM, Armonk, NY). All the results were considered statistically significant at the .05 level.

RESULTS

Subject Characteristics

Study cohort population characteristics are provided in Table 1, and PRM results are displayed in Table 2. No significant differences in BMI, height, and weight were observed between strata. Never-smokers were found to be significantly younger than GOLD 1 subjects. As expected, lung volumes, FEV1, and PRM metrics at both interval examinations were found to be dependent on GOLD. No significant correlations were observed between FEV1 and PRM classifications within the stratum and at individual examinations. In addition, no significant relationships were obtained between subject gender and GOLD status. Finally, change in PRM^{Normal}, PRM^{fSAD}, and PRM^{Emph} for each GOLD stratum was found to be insignificant over the 30-day interval (all cases with $P>.07$). The 95% confidence intervals in changes in PRM metrics over the 30-day interval are presented in the online supplemental Results.

HU Variability for Interval CT Scans

All subjects' generated histograms for Exp and Ins were similar to normal distributions. Of the 90 subjects, 90% were found to have histograms with a t-location scale probability distribution, whereas in the remaining 10%, histograms were generated with a logistic distribution. These functions, like Student t, are symmetric about its mean value with a leptokurtic shape. We found that the means of the fitted distribution functions over all 90 subjects and breath-holds were negligibly different from the expected value of 0.

Our procedure for applying a linear regression to the principal component analysis-transformed serial CT data at inspiration and expiration was found to generate consistent results irrespective of GOLD status or ventilation. To illustrate this point, Figure 2 presents voxel HU scatter plots at inspiration and expiration of subjects diagnosed with GOLD 1 and GOLD 3 COPD. Although the location of the voxel distribution varied between the 2 cases, the regression fits (red lines in Figure 2) generated consistent slopes and Y-intercepts (inserts in Figure 2). This consistency was also verified over the entire population (online supplemental Results). As expected, IMV was found to decrease with increasing COPD severity (Figure 2). As the disease progresses, lung parenchymal density approaches ambient air.

To address limitations in HU variability at low density, IMV was modeled as a function of HU. We assumed that the serial pair of modes was an adequate approximation of the centroid (center of mass) for the density distribution observed in the scatter plots. The dependence of IMV on HU values is particularly clear in Figure 3, in which IMV values show a nonlinear drop in value with decreasing HU average of modes. To strengthen the fit of IMV model to the data, expiration (blue markers in Figure 3) and inspiration (red markers in Figure 3) data were pooled. The optimal parameters obtained for the IMV model were $V = 66.5$ and $r = -0.014$ with a goodness of fit of NRMSE = 0.87 as defined as the normalized root mean squared error (online supplemental Results).

Application of Correction Approach on Follow-up PRM

Figure 4 shows representative sections from subjects diagnosed as GOLD 1 and 3 COPD at baseline; these are the same subjects shown in Figure 2. In the first case, PRM^{fSAD} was found to drop by 2.8% at follow-up (Figure 4, top row fourth column) with no correction strategy implemented. Processing the serial CT scans using the strategy outlined in Figure 1 generated a corrected PRM^{fSAD} that resulted in a drop from baseline of only 0.3% (Figure 4, top row fifth column). Consequent to the low relative volume of emphysema as determined by PRM, negligible benefits were observed when correcting the follow-up PRM. Nevertheless, the effect of our correction is evident locally near the apex of the left lung for subject GOLD 1 (Figure 4 top row black arrow fourth and fifth columns). The subject diagnosed with GOLD 3 COPD showed a decrease in emphysematous lung voxels, going from 28.7% at 0-day to 26.4% at the 30-day CT scan, yielding a short-term drop of 2.3% in emphysema (Figure 4, bottom row fourth column). Applying our correction scheme on the follow-up PRM resulted in a difference in PRM^{Emph} of only 0.3%. A similar result was observed for PRM^{fSAD}, with a change of 2.8% for the original PRM^{fSAD} and 0.4% for the corrected PRM^{fSAD}. Upon closer inspection of the bottom right lung, PRM voxel classification (w/o correction) at follow-up varies substantially from the baseline PRM for this subject (Figure 4 bottom row blue arrows).

The distributions of PRM^{Norm}, PRM^{fSAD}, and PRM^{Emph} separated by GOLD, for both uncorrected and corrected models, are presented in Figure 5. As expected, never-smokers and GOLD 0 subjects showed prevalence of voxels classified as PRM^{Norm} with an interquartile range (IQR) of <2.4% for PRM^{Norm} and PRM^{fSAD} in both never-smokers and GOLD 0. As the disease severity increased toward more fSAD and emphysema, erroneous shifts in PRM classifications were more prevalent. We observed maximum classification variability on GOLD 2 and GOLD 3 subjects, yielding an IQR of >8.4% for PRM^{Norm} and PRM^{fSAD}. In the case of GOLD 3, we found the largest PRM^{Emph} mismatch with IQR = 2.3%. Applying our correction strategy relieves the level of noise in PRM classification to more uniform PRM distributions, with an IQR of <1.7% in all cases.

To assess the agreement in PRM classification, overall percent agreement scores were determined between baseline PRM and the 30-day CT scans either with or without corrections of PRM voxel classification. Figure 6 displays the distribution of the percentage agreement (%agreement) scores for follow-up PRM results over GOLD stages. We found significant differences between the uncorrected and corrected models, both globally ($P< .0001$) and within GOLD ($P< .01$ in all cases). Uncorrected follow-up PRM values were found to generate %agreement that significantly varied with increasing GOLD ($P< .0001$). This trend was not observed for %agreement using the corrected follow-up PRM (Figure 6).

DISCUSSION

We propose a strategy that addresses voxel-level measurement variability in serially aligned inspiratory–expiratory paired CT scans that affect voxel classification by PRM. CT data acquired at inspiration and expiration at a 30-day interval were used to quantify the variability of HU values owing to system noise and alignment imperfections during postprocessing. After verifying that the change in voxel HU values from baseline and at

follow-up preserves the properties of a normal distribution, we defined an IMV that adjusts as a function of HU measurements. Our correction strategy allows voxel-level classification shifts to occur only when changes in voxel HU values exceed IMV(<HU>). We found that voxel-level agreement between uncorrected follow-up and baseline PRM data worsened with disease severity. Our strategy of accounting for system noise diminished this trend in agreement between longitudinal PRM data.

Even with detailed spatial information present in CT imaging, the standard approach for assessing and monitoring disease by this modality has often been limited to calculating whole-lung or large tissue (ie, lobes) measures. Typically based on summary statistics (ie, mean), such as mean airway wall measurements (17) or percent air trapping using only an expiratory CT scan acquisition, variability in longitudinal CT scans has led to erroneous conclusions (18). For example, recently, Smith et al. (19) showed in a comparison of spatially matched airway segments of nonsmokers with a COPD population, that COPD subjects were found to have thinner walls, contrary to the data relying on a more lumped approach. The same whole-lung approach has also been applied to spatially aligned data, such as PRM, where individual classifications are presented as percentages of the entire lung volume (4, 20). In response to the findings from McDonough et al. where they postulated that SAD is an intermediate step toward emphysema (21); Boes et al proposed an approach for assessing voxel-level changes in PRM classification, which was referred to as "voxel-based tracking" (5). One-year interval paired CT data from SPIROMICS were spatially aligned to a single geometric frame, in this case the baseline expiration CT scan. As described in this study, each voxel in the lung parenchyma consisted on 2 temporally resolved PRM classifications. In a subject diagnosed with GOLD 2 COPD and found to have a decline in absolute FEV1 from 2.34 L to 2.12 L over 1 year, 48% of all voxels identified by PRM as emphysema at follow-up were fSAD at baseline. The ability to monitor COPD at the voxel-level could provide clinicians with a means to monitor local progression. Although promising, the effect of measurement variability on PRM classification was evident in the whole-lung measures analyzed in that study, where PRM^{Emph} was found to decrease in a small number of subjects over the 1-year period (5), similar to our finding in Figure 4 for the GOLD 3 subject. Our observations concluded that fSAD measurements from uncorrected follow-up PRM, on average, showed moderate agreement to baseline PRM (Figure 5), even within the relatively short time frame of 30 days. The deterioration in PRM^{fSAD} agreement with increasing GOLD status is attributed to the large number of voxels with HU values near the −950 HU and −856 HU thresholds for GOLD 1–3 subjects. Even a small deviation in the HU value between longitudinal CT scans (<IMV) could result in a shift in PRM classification. For example, a voxel may easily shift over −856 HU from normal to/from fSAD (Norm ↔ fSAD) or over −950 HU from fSAD to/from Emph (fSAD ↔ Emph). Through the proposed approach, we have provided a strategy to mitigate erroneous shifts in voxel-level PRM classification when analyzing longitudinal data by PRM.

Limitations in our approach deserve further attention. Although CT scans from 90 subjects were available for development of the approach, the composition of the population significantly varied across GOLD (Table 1). Quantitative CT values are highly dependent on scanner vendor, scanner type, acquisition parameters (eg, kV, mA, and FOV) and reconstruction kernels. In our study population, a subset of subjects underwent serial CT

scans on different scanner types and acquired at different FOVs, with both limitations producing negligible differences in %agreement (details in online supplemental Results). In addition, we did not correct for inadequate ventilation at serial CT examinations (discussion in online supplemental Results). The effect of HU variability, consequent of CT acquisition, processing, and inadequate ventilation, on PRM quantification, has been previously reported, and techniques for alleviating their effect has been discussed (6). Although this approach does not address all errors associated with evaluating serial CT scans, our strategy for correcting PRM classification shifts is highly adaptable, allowing additional techniques that resolve more specific sources of error to be easily integrated in our workflow.

Consequent to the impact of COPD to health systems worldwide, extensive research is being devoted to the development and evaluation of novel biomarkers. PRM has been shown in multiple studies to serve as an objective and quantitative measure of disease. Large-scale multicenter observational studies such as COPDGene and SPIROMICS provided temporally re-solved HRCT to evaluate metrics, such as PRM, for monitoring COPD progression and, ideally, for therapeutic response assessment. Our methodology for correcting shifts in PRM classification due to variability in longitudinal HRCT scans may improve the clinical management of patients through more accurate monitoring of COPD subtypes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Abbreviations

SAD	Small airways disease
COPD	chronic obstructive pulmonary disease
PRM	parametric response mapping
CT	computed tomography
HRCT	high-resolution computed tomography
fSAD	functional SAD
HU	Hounsfield unit
FEV1	forced expiratory volume at 1 second

FOV   field of view

BMI   body mass index

IMV   index of measurement variability

IQR   interquartile range

References

1. Bousquet J, Dahl R, Khaltaev N. Global alliance against chronic respiratory diseases. Eur Respir J. 2007; 29(2):233–239. [PubMed: 17264322]

2. Mannino DM, Buist AS. Global burden of COPD: risk factors, prevalence, and future trends. Lancet. 2007; 370(9589):765–773. [PubMed: 17765526]

3. Han MK. Clinical correlations of computed tomography imaging in chronic obstructive pulmonary disease. Ann Am Thorac Soc. 2013; 10(Suppl):S131–S137. [PubMed: 24313763]

4. Galbán CJ, Han MK, Boes JL, Chughtai KA, Meyer CR, Johnson TD, Galbán S, Rehemtulla A, Kazerooni EA, Martinez FJ, Ross BD. Computed tomography-based biomarker provides unique signature for diagnosis of COPD phenotypes and disease progression. Nat Med. 2012; 18(11):1711–1715. [PubMed: 23042237]

5. Boes JL, Hoff BA, Bule M, Johnson TD, Rehemtulla A, Chamberlain R, Hoffman EA, Kazerooni EA, Martinez FJ, Han MK, Ross BD, Galbán CJ. Parametric response mapping monitors temporal changes on lung CT scans in the subpopulations and intermediate outcome measures in COPD Study (SPIROMICS). Acad Radiol. 2015; 22(2):186–194. [PubMed: 25442794]

6. Boes JL, Bule M, Hoff BA, Chamberlain R, Lynch DA, Stojanovska J, Martinez FJ, Han MK, Kazerooni EA, Ross BD, Galbán CJ. The impact of sources of variability on parametric response mapping of lung CT scans. Tomography. 2015; 1(1):69–77. [PubMed: 26568983]

7. Bhatt SP, Soler X, Wang X, Murray S, Anzueto AR, Beaty TH, Boriek AM, Casa-buri R, Criner GJ, Diaz AA, Dransfield MT, Curran-Everett D, Galbán CJ, Hoffman EA, Hogg JC, Kazerooni EA, Kim V, Kinney GL, Lagstein A, Lynch DA, Make BJ, Martinez FJ, Ramsdell JW, Reddy R, Ross BD, Rossiter HB, Steiner RM, Strand MJ, van Beek EJ, Wan ES, Washko GR, Wells JM, Wendt CH, Wise RA, Silverman EK, Crapo JD, Bowler RP, Han MK, COPDGene Investigators. 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]

8. Galbán CJ, Boes JL, Bule M, Kitko CL, Couriel DR, Johnson TD, Lama V, Telenga ED, van den Berge M, Rehemtulla A, Kazerooni EA, Ponkowski MJ, Ross BD, Yanik GA. Parametric response mapping as an indicator of bronchiolitis obliterans syndrome after hematopoietic stem cell transplantation. Biol Blood Marrow Transplant. 2014; 20(10):1592–1598. [PubMed: 24954547]

9. Verleden SE, Vos R, Vandermeulen E, Ruttens D, Bellon H, Heigl T, Van Raemdonck DE, Verleden GM, Lama V, Ross BD, Galbán CJ, Vanaudenaerde BM. Parametric response mapping of bronchiolitis obliterans syndrome progression after lung transplantation. Am J Transplant. 2016; 16(11):3262–3269. [PubMed: 27367568]

10. Belloli EA, Degtiar I, Wang X, Yanik GA, Stuckey LJ, Verleden SE, Kazerooni EA, Ross BD, Murray S, Galbán CJ, Lama VN. Parametric response mapping as an imaging biomarker in lung transplant recipients. Am J Respir Crit Care Med. 2017; 195(7):942–952. [PubMed: 27779421]

11. Couper D, LaVange LM, Han M, Barr RG, Bleecker E, Hoffman EA, Kanner R, Kleerup E, Martinez FJ, Woodruff PG, Rennard S, SPIROMICS Research Group. Design of the subpopulations and intermediate outcomes in COPD study (SPIROMICS). Thorax. 2014; 69(5):491–494.

12. Rabe KF, Hurd S, Anzueto A, Barnes PJ, Buist SA, Calverley P, Fukuchi Y, Jenkins C, Rodriguez-Roisin R, van Weel C, Zielinski J, Global Initiative for Chronic Obstructive Lung Disease. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. Am J Respir Crit Care Med. 2007; 176(6):532–555. [PubMed: 17507545]

13. Sieren JP, Newell JD Jr, Barr RG, Bleecker ER, Burnette N, Carretta EE, Couper D, Goldin J, Guo J, Han MK, Hansel NN, Kanner RE, Kazerooni EA, Martinez FJ, Rennard S, Woodruff PG, Hoffman EA, SPIROMICS Research Group. SPIROMICS protocol for multicenter quantitative computed tomography to phenotype the lungs. Am J Respir Crit Care Med. 2016; 194(7):794–806. [PubMed: 27482984]

14. Sieren JP, Newell JD, Judy PF, Lynch DA, Chan KS, Guo J, Hoffman EA. Reference standard and statistical model for intersite and temporal comparisons of CT attenuation in a multicenter quantitative lung study. Med Phys. 2012; 39(9):5757–5767. [PubMed: 22957640]

15. Hamstra DA, Galbán CJ, Meyer CR, Johnson TD, Sundgren PC, Tsien C, Lawrence TS, Junck L, Ross DJ, Rehemtulla A, Ross BD, Chenevert TL. Functional diffusion map as an early imaging biomarker for high-grade glioma: correlation with conventional radiologic response and overall survival. J Clin Oncol. 2008; 26(20):3387–3394. [PubMed: 18541899]

16. Galbán CJ, Chenevert TL, Meyer CR, Tsien C, Lawrence TS, Hamstra DA, Junck L, Sundgren PC, Johnson TD, Ross DJ, Rehemtulla A, Ross BD. The parametric response map is an imaging biomarker for early cancer treatment outcome. Nat Med. 2009; 15(5):572–576. [PubMed: 19377487]

17. Nambu A, Zach J, Schroeder J, Jin G, Kim SS, Kim YI, Schnell C, Bowler R, Lynch DA. Quantitative computed tomography measurements to evaluate airway disease in chronic obstructive pulmonary disease: relationship to physiological measurements, clinical index and visual assessment of airway disease. Eur J Radiol. 2016; 85(11):2144–2151. [PubMed: 27776670]

18. Newman KB, Lynch DA, Newman LS, Ellegood D, Newell JD Jr. Quantitative computed tomography detects air trapping due to asthma. Chest. 1994; 106(1):105–109. [PubMed: 8020254]

19. Smith BM, Hoffman EA, Rabinowitz D, Bleecker E, Christenson S, Couper D, Donohue KM, Han MK, Hansel NN, Kanner RE, Kleerup E, Rennard S, Barr RG. Comparison of spatially matched airways reveals thinner airway walls in COPD. The Multi-Ethnic Study of Atherosclerosis (MESA) COPD Study and the Subpopulations and Intermediate Outcomes in COPD Study (SPIROMICS). Thorax. 2014; 69(11):987–996. [PubMed: 24928812]

20. Boudewijn IM, Postma DS, Telenga ED, Ten Hacken NH, Timens W, Oudkerk M, Ross BD, Galbán CJ, van den Berge M. Effects of ageing and smoking on pulmonary computed tomography scans using parametric response mapping. Eur Respir J. 2015; 46(4):1193–1196. [PubMed: 26113678]

21. McDonough JE, Yuan R, Suzuki M, Seyednejad N, Elliott WM, Sanchez PG, Wright AC, Gefter WB, Litzky L, Coxson HO, Paré PD, Sin DD, Pierce RA, Woods JC, McWilliams AM, Mayo JR, Lam SC, Cooper JD, Hogg JC. Small-airway obstruction and emphysema in chronic obstructive pulmonary disease. N Engl J Med. 2011; 365(17):1567–1575. [PubMed: 22029978]
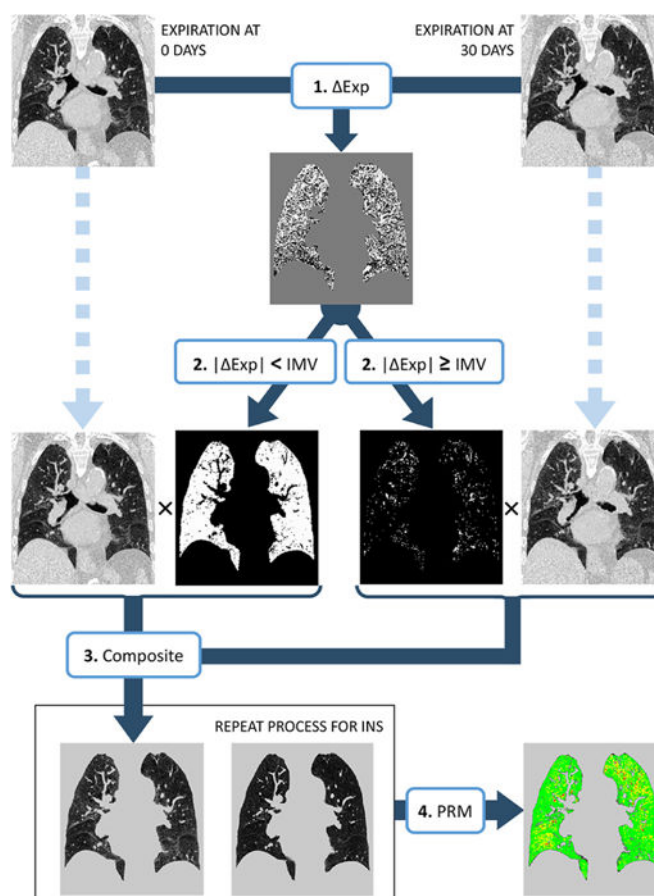
Figure 1.
Workflow summarizing the correction of erroneous PRM classifications.
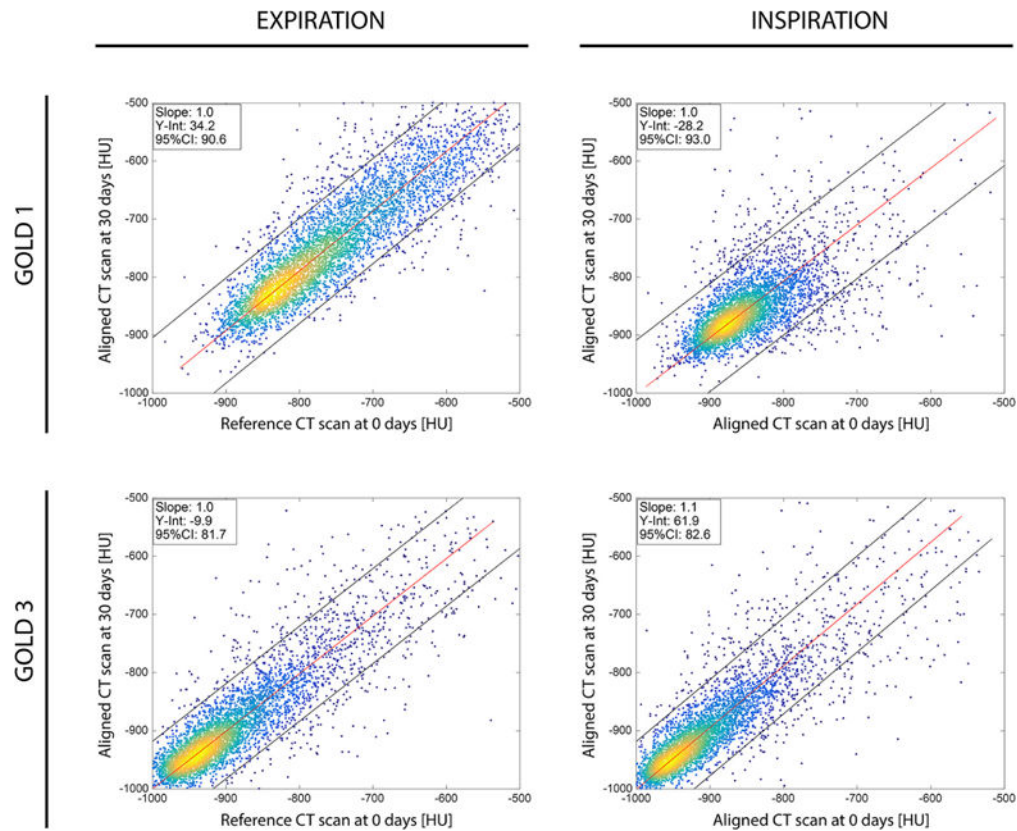
Figure 2.
Density scatter plots of voxels with interval HU values acquired at expiration (left) and inspiration (right) are presented for representative GOLD 1 (top) and GOLD 3 (bottom) subjects. The regression lines (red line) and 95% confidence intervals (index of measurement variability [IMV]; black lines) from the fit of the data transformed using the principal component analysis are included in the plots. Values of the slope and Y-intercept of the fit are shown at the upper left corner of each plot, jointly with the value of the IMV.
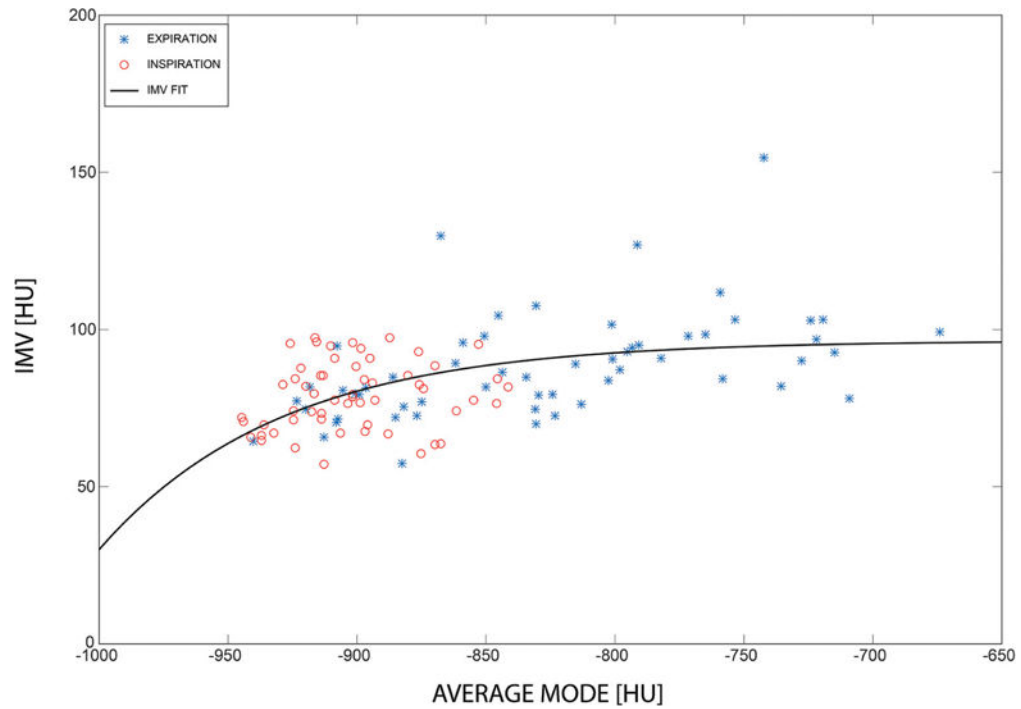
Figure 3.
Presented is a plot of IMV as a function of the Hounsfield unit (HU) modes, with values averaged over 0 and 30 days, for each patient. A nearly sigmoid-shaped model was nonlinearly fitted to the data from subjects with relative changes in lung volume below 10%. Data for serial inspiration and expiration CT scans were pooled for this analysis. The optimal model for the presented data was: $IMV(x) = 66.5 \, [1 - e^{-0.014 \, (x + 1000)}] + 30$.
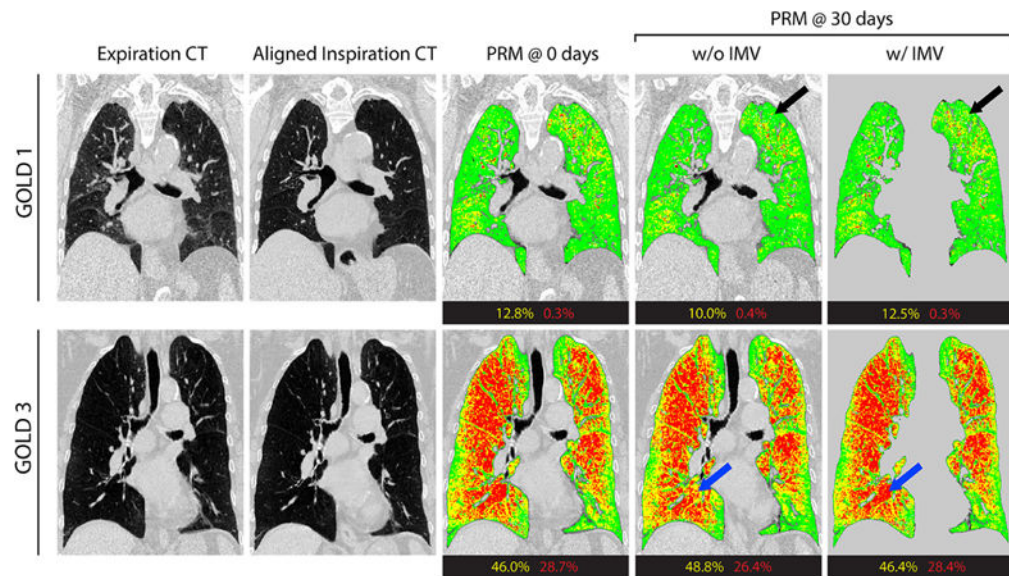
Figure 4.
Presented are representative coronal sections from baseline expiration and aligned inspiration computed tomography (CT) scans, and parametric response mapping (PRM) images generated using baseline scans and 30-d follow-up scans without and with IMV correction for the GOLD 1 and 3 subjects in Figure 2. Individual voxels are color-coded per PRM as follows: green for PRM^{Normal}, yellow for PRM^{fSAD}, and red for PRM^{Emph}. The relative volume of PRM^{fSAD} (yellow) and PRM^{Emph} (red) are included below each PRM image. Regions of discrepancy in PRM classification are indicated in the upper left lung (black arrows) and lower right lung (blue arrows) of the GOLD 1 and GOLD 3 subjects, respectively.
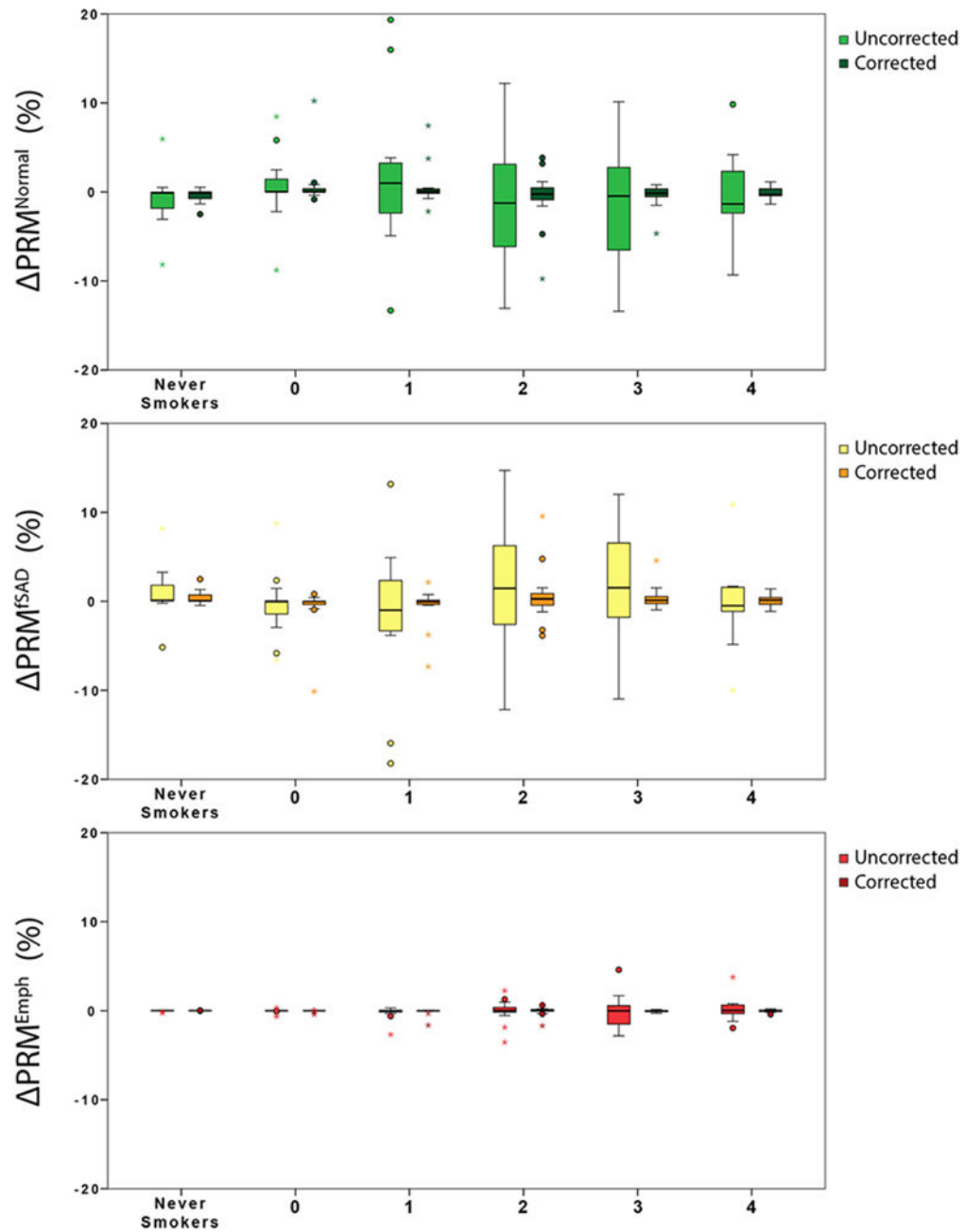
Figure 5.
The difference in whole-lung serial percent volume of PRM^{Normal} (top), PRM^{fSAD} (center), and PRM^{Emph} (bottom) over GOLD with and without correction. Data are presented as box-plots with center line at the median, box limits at the 25th percentile and 75th percentile, and whiskers at the 5th percentile and 95th percentile.

Figure 6.
Percent agreement of serial voxel-level PRM classifications over GOLD with and without
correction. Data are presented as boxplots with center line at the median, box limits at the
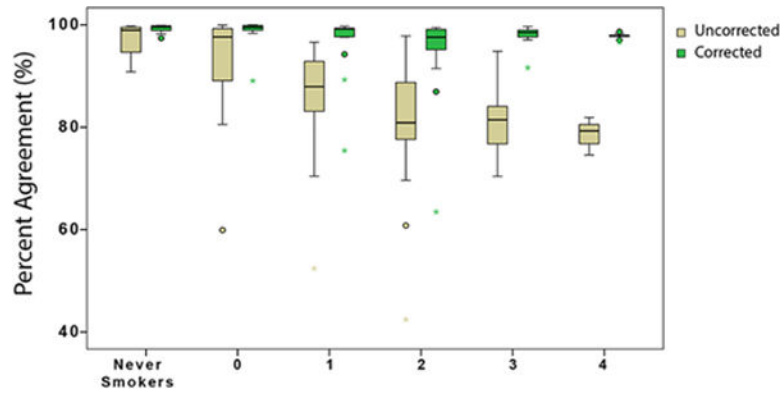25th percentile and 75th percentile, and whiskers at the 5th percentile and 95th percentile.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Subject Characteristics

Parameter	Never-Smokers	GOLD 0	GOLD 1	GOLD 2	GOLD 3	GOLD 4
Number	11	17	15	18	20	9
Gender (M/F)	4/7	10/7	13/2	13/5	13/7	5/4
Age (years)	55 (7)	55 (8)	67 (8)	64 (8)	64 (9)	61 (9)
Height (cm)	170 (12)	169 (9)	172 (6)	172 (10)	172 (10)	169 (13)
Weight (kg)	79 (20)	82 (20)	79 (15)	84 (18)	81 (24)	78 (20)
BMI (kg/cm^2)	27 (4)	29 (5)	27 (5)	28 (6)	27 (6)	27 (5)
FEV1 (% predicted)						
0 days	110 (7)	98 (12)	92 (10)	65 (9)	41 (6)	26 (4)
30 days	107 (9)	92 (12)	86 (13)	58 (12)	36 (8)	24 (8)
Exp volume (L)						
0 days	2.39 (0.61)	2.81 (0.82)	3.36 (1.08)	3.87 (1.11)	4.57 (1.22)	5.09 (0.99)
30 days	2.49 (0.67)	2.68 (0.55)	3.17 (0.71)	3.81 (0.85)	4.53 (1.00)	5.11 (0.97)
Ins volume (L)						
0 days	5.93 (1.68)	5.59 (0.95)	5.96 (1.34)	6.09 (1.45)	6.31 (1.39)	6.53 (1.31)
30 days	5.76 (1.61)	5.49 (1.05)	5.88 (1.24)	6.28 (1.32)	6.38 (1.50)	6.55 (1.33)

Subject characteristics by GOLD stage. Values are mean (standard deviation) (except for gender). Characteristics not temporally disaggregated were recorded at 0 day.

Abbreviations: BMI, body mass index; FEV1, forced expiratory volume in 1 second (percentage of predicted); Ins and Exp, inspiration and expiration, respectively.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 2

PRM Results

PRM	Never-Smokers	GOLD 0	GOLD 1	GOLD 2	GOLD 3	GOLD 4
Normal						
0 days	98.31 (3.1)	91.25 (17.1)	81.33 (17.9)	69.30 (22.6)	53.63 (24.0)	36.13 (12.9)
30 days	97.50 (3.1)	93.93 (10.5)	85.43 (10.8)	70.03 (18.8)	52.06 (23.5)	36.12 (15.5)
fSAD						
0 days	1.38 (2.7)	8.32 (16.7)	16.08 (15.4)	23.06 (19.0)	33.62 (13.9)	46.77 (7.8)
30 days	2.29 (3.1)	5.75 (10.2)	12.08 (7.7)	22.42 (13.8)	35.35 (14.4)	46.56 (8.1)
Emphysema						
0 days	0.06 (0.1)	0.18 (0.4)	1.96 (4.2)	6.84 (13.8)	12.26 (13.3)	16.78 (10.7)
30 days	0.04 (0.03)	0.13 (0.3)	1.74 (3.6)	6.80 (13.8)	12.02 (12.8)	16.97 (10.5)

Percent lung volumes of PRM^{Normal}, PRM^{fSAD}, and PRM^{Emph} over GOLD stages. Values are presented as mean (standard deviation). Spatial alignment of data was not corrected for insufficient ventilation between time points.

Abbreviation: fSAD, functional small airways disease.