

RESEARCH

Open Access

Epidemiology, genetics, and subtyping of preserved ratio impaired spirometry (PRISm) in COPDGene

Emily S Wan^{1,2*}, Peter J Castaldi¹, Michael H Cho^{1,2}, John E Hokanson³, Elizabeth A Regan⁴, Barry J Make⁴, Terri H Beaty⁵, MeiLan K Han⁶, Jeffrey L Curtis⁷, Douglas Curran-Everett^{8,9}, David A Lynch⁴, Dawn L DeMeo^{1,2}, James D Crapo⁴, Edwin K Silverman^{1,2} and The COPDGene Investigators

Abstract

Background: Preserved Ratio Impaired Spirometry (PRISm), defined as a reduced FEV₁ in the setting of a preserved FEV₁/FVC ratio, is highly prevalent and is associated with increased respiratory symptoms, systemic inflammation, and mortality. Studies investigating quantitative chest tomographic features, genetic associations, and subtypes in PRISm subjects have not been reported.

Methods: Data from current and former smokers enrolled in COPDGene (n = 10,192), an observational, cross-sectional study which recruited subjects aged 45–80 with ≥10 pack years of smoking, were analyzed. To identify epidemiological and radiographic predictors of PRISm, we performed univariate and multivariate analyses comparing PRISm subjects both to control subjects with normal spirometry and to subjects with COPD. To investigate common genetic predictors of PRISm, we performed a genome-wide association study (GWAS). To explore potential subgroups within PRISm, we performed unsupervised k-means clustering.

Results: The prevalence of PRISm in COPDGene is 12.3%. Increased dyspnea, reduced 6-minute walk distance, increased percent emphysema and decreased total lung capacity, as well as increased segmental bronchial wall area percentage were significant predictors (p-value <0.05) of PRISm status when compared to control subjects in multivariate models. Although no common genetic variants were identified on GWAS testing, a significant association with Klinefelter's syndrome (47XXY) was observed (p-value < 0.001). Subgroups identified through k-means clustering include a putative "COPD-subtype", "Restrictive-subtype", and a highly symptomatic "Metabolic-subtype".

Conclusions: PRISm subjects are clinically and genetically heterogeneous. Future investigations into the pathophysiological mechanisms behind and potential treatment options for subgroups within PRISm are warranted.

Trial registration: Clinicaltrials.gov Identifier: NCT000608764.

Keywords: Spirometry, Restriction, Lung diseases, Smoking

Background

Since its inception in the mid-19th century [1], spirometry has become an accepted tool in the diagnosis and staging of obstructive lung diseases (defined as the disproportionate reduction in the forced expiratory volume in the first second (FEV₁) relative to the forced vital

capacity (FVC)). However, subjects with substantial, *proportionate* impairments in FEV₁ and FVC resulting in a preserved FEV₁/FVC ratio have remained a relatively understudied group. Approximately 1 out of every 8 subjects in the general population has Preserved Ratio Impaired Spirometry (PRISm); this pattern has alternatively been referred to as "unclassified", "non-specific", or "restrictive" spirometry, with the latter term being the most widely accepted. It should be noted, however, that the predictive value of "restrictive spirometry" for true restriction, as defined by a reduced total lung capacity, is

* Correspondence: emily.wan@channing.harvard.edu

¹Channing Division of Network Medicine, Brigham and Women's Hospital, 181 Longwood Avenue, 4th floor, Boston, MA 2115, USA

²Division of Pulmonary and Critical Care Medicine, Brigham and Women's Hospital, Boston, MA, USA

Full list of author information is available at the end of the article

poor [2–4]. In addition, although restriction is often associated with interstitial lung diseases, many PRISm subjects do *not* have evidence for interstitial lung diseases on radiographic evaluation [5]. Less commonly used terms, such as “non-specific” [6] or “unclassified” spirometry [7,8], do not make inferences regarding the etiology of the spirometric abnormalities, but are generally uninformative.

The cross-sectional prevalence of PRISm has been estimated to be between 6.6%–17.6% [9–16] worldwide. While local and regional variability in the prevalence of PRISm exists [13,15], these estimates remain stable regardless of whether the Global Initiative for Obstructive Lung Disease [17] (GOLD) or lower limits of normal (LLN) diagnostic criteria are utilized [8,14,16]. Subjects with PRISm have increased morbidity [8–10,12,15,16,18] and mortality [6,9,16,18]. They report increased respiratory symptoms [9,16,19], decreased exercise capacity [12,19], more difficulty with the activities of daily living [12,15], and have evidence of increased systemic inflammation [20]. Additional trends which have emerged include associations between PRISm and increased body mass index (BMI) [6,9,10,15], diabetes mellitus [8,12,13,18,21], cardiovascular disease [9,13,18,22], and cigarette smoke exposure [6,10,13,15,18]. While these summary statistics among all PRISm subjects are useful, they fail to capture the significant heterogeneity present within this group; for example, while the mean BMI of this cohort is typically higher than that of the general population, the range of BMI observed can include frankly cachectic subjects [8,10,12,13,15,18]. In our previous analysis of PRISm subjects among the first 2,500 subjects from COPD-Gene [8], we hypothesized that this heterogeneity reflected the multitude of potential underlying causes for this spirometric pattern. Using data from current and former smokers enrolled in the full COPDGene cohort, we now seek to examine the following hypotheses:

Aim 1: We hypothesize that a distinct set of predictors are associated with PRISm status and examine the epidemiological, functional, and radiographic predictors of PRISm status relative to control and COPD subjects.

Aim 2: We hypothesize that genetic variants may contribute to the development of PRISm among current and former smokers and explore the associations between common genetic variants and PRISm status relative to control subjects.

Aim 3: We hypothesize that subgroups exist within the PRISm cohort and explore the utility of unbiased machine learning approaches in identifying potentially pathobiologically distinct groups within PRISm.

Materials and methods

Study population

All subjects were participants in COPDGene (ClinicalTrials.gov Identifier NCT000608764); enrollment and exclusion

criteria have been previously described [23]. COPDGene is a cross-sectional, observational study which enrolled self-identified non-Hispanic white (NHW) or African American (AA) current or former smokers aged 45–80 years with ≥ 10 pack-years of smoking. Institutional review board approval was obtained at each of the 21 participating clinical centers (please see Additional file 1–Additional Methods section for the names of the approving IRB offices); all subjects provided written informed consent. Subjects completed questionnaires, pre- and post-bronchodilator spirometry, 6-minute walk test, and volumetric chest computed tomography (CT) at full inspiration and expiration. All analyses were conducted using the COPDGene phenotype dataset released September 19, 2012.

Variable definitions

Percent predicted values and lower limits of normal (LLN) were calculated using post-bronchodilator spirometric values [24]. Fixed threshold-defined groups were as follows: PRISm subjects had an $FEV_1 < 80\%$ predicted with an $FEV_1/FVC \geq 0.7$, control subjects had an $FEV_1 \geq 80\%$ with an $FEV_1/FVC \geq 0.7$, and COPD subjects had an $FEV_1 < 80\%$ predicted with an $FEV_1/FVC < 0.7$. The distribution of spirometry by FEV_1 and FEV_1/FVC in the COPDGene cohort is illustrated in Figure 1. LLN-defined cohorts were defined as follows: LLN-PRISm subjects had $FEV_1 < LLN$ with an $FEV_1/FVC \geq LLN$, LLN-controls had $FEV_1 \geq LLN$ with an $FEV_1/FVC \geq LLN$, while LLN-COPD subjects had $FEV_1 < LLN$ and $FEV_1/FVC < LLN$. The distribution of spirometry and delineation of LLN-defined populations are illustrated in Additional file 1: Figure S1. Additional variable definitions are available in detail (see Additional file 1).

Aim 1: Epidemiological analysis

Univariate comparisons between PRISm subjects and control or COPD subjects were made using a Student's t-test or Wilcoxon rank sum test for normal and non-normally distributed continuous variables, respectively, while Chi-square or Fisher's exact tests were used for discrete variables. All nominally significant variables ($P_{univariate} < 0.05$), except lung function and medication use variables, were considered candidate independent variables for multivariate regression. Logistic regression using automated stepwise selection with binary PRISm status as the dependent variable was performed using SAS (v 9.3, Cary, NC); a significance level of 0.1 was specified for entry into the model and independent variables with a p-value < 0.05 were retained in the final model. Additional details regarding stepwise selection are outlined in the Additional file 1 – Methods section. Non-significant candidate independent variables were

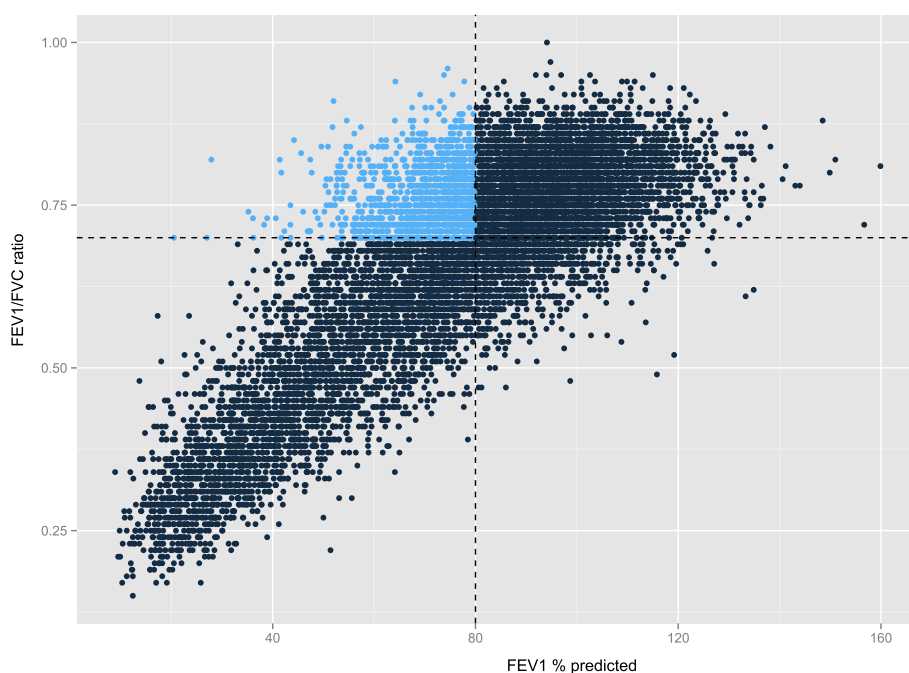


Figure 1 Distribution of spirometry in COPDGene. Legend: Forced expiratory volume in the first second (FEV_1) is plotted on the x-axis while FEV_1/FVC ratio is plotted on the y-axis. Dashed lines represent fixed-threshold criteria used to delineate Preserved Ratio Impaired Spirometry (PRISm) subjects (highlighted in blue-upper left quadrant), control (upper right quadrant), and COPD (lower left quadrant) subjects.

tested as confounders and were retained if >10% change in the effect estimate was observed.

Aim 2: Genetic analysis

Genome-wide single nucleotide polymorphism (SNP) genotyping data were obtained on the Illumina (San Diego, CA, USA) OmniExpress platform with additional genotypes imputed using MaCH [25] software and the 1000 Genomes [26] phased data (Hg19) for a total of ~14.9 million SNPs. Additional details regarding data processing and quality control are available (see Additional file 1). Genome-wide association testing for associations with binary PRISm status relative to control subjects was performed using an additive model, adjusted for age, sex, pack-years of smoking, BMI, current smoking status and principal components for genetic ancestry. Analyses were performed separately in non-Hispanic white and African American subjects using PLINK [27]; meta-analysis using standard error weighting was then performed using METAL software [28].

Aim 3: Unsupervised cluster analysis to identify potential subgroups within PRISm

Unsupervised k-means clustering analyses were conducted on the subset of PRISm subjects ($n_{\text{fixed threshold}} = 1,135$ and $n_{\text{LLN}} = 978$) with complete data for six empirically chosen key input variables: $TLC_{CT}\%$ predicted, $FEV_1\%$ predicted, FEV_1/FVC ratio, percent emphysema

($\%LAA-950_{\text{insp}}$), BMI, and segmental wall area percent. Normalized mutual information (NMI) using a five-fold cross-validation strategy was used to determine the optimal number of clusters [29]. To assess whether differences in the number or types of subgroups identified differed by race and to reduce the impact of population stratification, clustering and candidate gene testing were performed separately in NHW and AA subjects using R (2.15.0) [30]. Additional details regarding the cluster analysis are available in the Additional file 1.

Results

Aim 1: Epidemiological, functional, and radiographic predictors of PRISm status

Among the current and former smokers enrolled in the COPDGene cohort ($n = 10,192$), the mean age was 59.6 years, the mean pack-years smoked was 44.2, and mean body mass index (BMI) was 28.8; 53.5% of the cohort was male, 33.4% were African American, and 53.1% were current smokers. By fixed-threshold criteria, PRISm subjects account for 12.3% ($n = 1,257$) of the COPDGene cohort (Figure 1). Univariate comparisons between PRISm subjects and 1) control subjects and 2) COPD subjects are summarized in Table 1. PRISm subjects have the highest proportion of females and current smokers, increased mean body mass index (BMI) and decreased mean total lung capacity, as well as an increased prevalence of diabetes mellitus relative

Table 1 Characteristics of Preserved Ratio Impaired Spirometry (PRISm), smoking control, and COPD subjects in COPDGene using fixed threshold criteria

| | Control subjects | PRISm | COPD |
|--|------------------|----------------|-----------------|
| n | 4388 | 1257 | 3690 |
| Age | 56.7 (8.4)* | 57.2 (8.2) | 63.4 (8.5)* |
| Sex (% Male) | 52.9* | 46.1 | 55.6* |
| African American | 41.2 | 43.1 | 22.7* |
| Current smoker | 59.7* | 63.8 | 40.7* |
| Pack-years | 37.2 (20.2)* | 42.7 (24.2) | 53.0 (27.5)* |
| Body Mass Index | 28.9 (5.8)* | 31.8 (7.3) | 28.1 (6.3)* |
| FEV ₁ % predicted | 97.5 (11.5)* | 70.2 (8.4) | 50.2 (18.0)* |
| FVC% predicted | 96.6 (11.9)* | 71.5 (9.1) | 76.3 (17.3)* |
| FEV ₁ /FVC | 0.79 (0.05)* | 0.77 (0.05) | 0.50 (0.13)* |
| Bronchodilator Responsiveness [†] | 10.0* | 13.7 | 36.6* |
| Total Lung Capacity _{CT} % predicted | 92.3 (14.6)* | 79.9 (13.5) | 101.5 (17.1)* |
| Segmental wall area percentage | 60.1 (2.9)* | 62.5 (3.1) | 62.9 (3.1)* |
| Percent emphysema (% LAA-950 _{insp}) | 2.0 (2.5)* | 1.4 (2.5) | 13.0 (12.8)* |
| Percent gas trapping (% LAA-856 _{exp}) | 11.0 (9.7) | 10.4 (9.1) | 39.2 (20.8)* |
| Pi10 | 3.65 (0.11)* | 3.73 (0.14) | 3.72 (0.14)* |
| 6 minute walk distance (feet) | 1491.5 (350.7)* | 1266.6 (366.9) | 1174.8 (397.0)* |
| MMRC [‡] Dyspnea score | 0.8 (1.2)* | 1.5 (1.5) | 2.1 (1.4)* |
| Resting O ₂ saturation | 97.1 (2.0)* | 96.5 (2.5) | 94.7 (3.6)* |
| Chronic bronchitis | 12.6* | 17.8 | 28.2* |
| Short acting beta-agonist use | 12.0* | 28.1 | 65.7* |
| Long acting beta-agonist use | 4.4* | 13.3 | 49.1* |
| Inhaled corticosteroid use | 5.5* | 16.4 | 51.2* |
| Oral corticosteroid use | 0.5* | 2 | 5.9* |
| Congestive heart failure | 1.3* | 4.6 | 5.4 |
| Coronary artery disease | 7.5* | 13.5 | 16.5* |
| Diabetes mellitus | 11.6* | 21.6 | 13.1* |
| Hypertension | 36.3* | 49.1 | 50.6 |
| Hyperlipidemia | 34.3* | 42.5 | 41.4 |
| History of blood clot | 2.8* | 5.2 | 5.7 |
| Peripheral vascular disease | 1.3* | 2.7 | 3.4 |
| History of stroke | 1.6* | 3.5 | 3.6 |
| Gastrointestinal reflux disease | 20.3* | 25.9 | 30.3* |
| History of compression fracture | 3.4* | 5.4 | 6.2 |
| Currently employed | 37.3* | 28.9 | 25.3* |
| Physician-diagnosed asthma | 11.4* | 21.1 | 24.6* |

Data are presented as mean (standard deviation) or percent. Preserved Ratio Impaired Spirometry (PRISm) defined as: FEV₁/FVC ≥ 0.7 & FEV₁ < 80% predicted. Control subjects defined as: FEV₁/FVC ≥ 0.7 & FEV₁ ≥ 80% predicted. Chronic Obstructive Pulmonary Disease (COPD) subjects defined as: FEV₁/FVC < 0.7 & FEV₁ < 80% predicted.

*Denotes univariate p-value < 0.05 when compared to PRISm subjects.

[†]Bronchodilator responsiveness considered present if the change in FEV₁ or FVC was >200 mL and ≥ 12% predicted following administration of short acting inhaled beta-agonist.

[‡]MMRC = modified Medical Research Council.

to both control and COPD subjects. PRISm subjects have significantly increased rates of respiratory medication use relative to smoking controls. However, within the PRISm

cohort, the correlation between respiratory medication use and a history of physician-diagnosed asthma or evidence of bronchodilator responsiveness (BDR) on

spirometry is poor (Additional file 1: Table S1); the simple kappa correlation between BDR and respiratory medication use ranged from 0.06 to 0.12. Thus, the majority of PRISm subjects with evidence of BDR do *not* report use of short- or long-acting beta agonists or inhaled steroids.

Because controversy regarding the use of fixed thresholds to define respiratory impairment exists, we repeated the above analyses using lower limit of normal criteria to define the PRISm, control, and COPD groups. The prevalence of lower limit of normal-defined PRISm (LLN-PRISm) is 10.6% (n = 1,082) of the final cohort; characteristics of these subjects relative to LLN – control and LLN – COPD subjects are summarized in Additional file 1: Table S2. LLN-PRISm subjects continue to demonstrate the highest mean BMI and lowest mean TLC% predicted, as well as the highest prevalence of diabetes mellitus; however, the enrichment of female subjects is no longer present.

The overlap between fixed threshold-defined PRISm and LLN-PRISm is illustrated in Additional file 1: Figure S2; 883 subjects are consistently identified as PRISm by both criteria (simple kappa coefficient = 0.72). Subjects with PRISm by LLN criteria only (n = 199) are significantly older, have increased emphysema and gas trapping as well as a lower FEV₁/FVC ratio relative to subjects identified as having PRISm by both fixed threshold and LLN criteria; 94.5% of these subjects (n = 188)

have Stage 2 COPD by Global Initiative for Obstructive Lung Disease (GOLD)¹¹ criteria. Of the subjects who have PRISm by fixed threshold criteria only (n = 374), 93.9% (n = 351) of these subjects are considered control subjects using LLN criteria. The reclassification of PRISm subjects by fixed-threshold and LLN-criteria is illustrated in Additional file 1: Figure S3.

Significant epidemiological predictors of PRISm status relative to control subjects in multivariate models for both fixed threshold and LLN-defined cohorts are presented in Table 2; none of the non-significant candidate independent variables were found to be confounders. The majority of risk factors identified were consistent regardless of whether fixed threshold or LLN criteria were used to define PRISm and control status and included increased cumulative pack-years, lower resting oxygen saturation, reduced 6-minute walk distance, increased MMRC dyspnea score, increased percent emphysema (after adjusting for TLC), decreased total lung capacity% predicted, increased segmental wall area percentage, and an increased prevalence of a history of peripheral vascular disease and physician-diagnosed asthma. Increased BMI and a history of diabetes mellitus were significant predictors only in the LLN cohort while female gender and increased age were significant only in the fixed threshold-defined cohort.

Multivariate models of PRISm status relative to both fixed-threshold and LLN-defined COPD are summarized

Table 2 Significant predictors of Preserved Ratio Impaired Spirometry (PRISm) status relative to control subjects in (a) fixed threshold-defined and (b) lower limit of normal (LLN)-defined cohorts in multivariate models

| Panel (a) – significant predictors in fixed threshold- defined cohorts | | Panel (b) – significant predictors in LLN defined cohorts | |
|--|---------------------|---|---------------------|
| Predictor | OR [95% CI] | Predictor | OR [95% CI] |
| Pack-years | 1.008 [1.005-1.012] | Pack-years | 1.008 [1.004-1.012] |
| Resting oxygen saturation | 0.907 [0.875-0.940] | Resting oxygen saturation | 0.931 [0.897-0.966] |
| 6 minute walk distance (per 100 feet) | 0.948 [0.926-0.971] | 6 minute walk distance (per 100 feet) | 0.953 [0.929-0.977] |
| MMRC Dyspnea score | 1.226 [1.153-1.304] | MMRC Dyspnea score | 1.192 [1.116-1.273] |
| Percent emphysema | 1.055 [1.020-1.091] | Percent emphysema | 1.061 [1.029-1.094] |
| Total lung capacity% predicted | 0.949 [0.943-0.955] | Total lung capacity% predicted | 0.948 [0.942-0.954] |
| Segmental wall area percent | 1.203 [1.168-1.238] | Segmental wall area percent | 1.190 [1.154-1.227] |
| Peripheral vascular disease | 1.777 [1.038-3.043] | Peripheral vascular disease | 2.038 [1.254-3.314] |
| Physician-diagnosed asthma | 1.391 [1.126-1.718] | Physician-diagnosed asthma | 1.485 [1.193-1.848] |
| Age | 1.025 [1.014-1.035] | Diabetes Mellitus | 1.372 [1.111-1.695] |
| Sex (male) | 0.705 [0.598-0.833] | African American race | 0.266 [0.216-0.328] |
| Hyperlipidemia | 1.216 [1.031-1.433] | Body mass index | 1.026 [1.012-1.040] |
| | | Current smoking | 1.620 [1.337-1.962] |
| | | Coronary artery disease | 1.370 [1.002-1.874] |

Variables tested but not retained in the final models:

Panel (a): body mass index, Pi10, current smoking, chronic bronchitis, bronchodilator responsiveness, congestive heart failure, coronary artery disease, diabetes, hypertension, history of blood clots, stroke, gastro-esophageal reflux disease, compression fractures, and current employment.

Panel (b): age, Pi10, chronic bronchitis, bronchodilator responsiveness, congestive heart failure, hypertension, hyperlipidemia, history of blood clots, stroke, gastro-esophageal reflux disease, compression fractures, and current employment.

No significant confounders (defined as causing >10% change in effect estimate) were found.

in Additional file 1: Table S3. Analogous to the comparison of PRISm and control subjects, the majority of predictors identified on multivariate modeling of PRISm vs. COPD subjects were consistent regardless of whether fixed-threshold or LLN criteria were used. Increased BMI relative to COPD subjects was consistently identified as a predictor of PRISm status, as was decreased bronchodilator responsiveness. Radiographic differences, such as decreased measurements of emphysema, gas trapping, TLC, and segmental wall area thickness were also among the robustly identified predictors of PRISm relative to COPD subjects.

Aim 2: Genetic associations between common genetic variants and PRISm status

During quality control of the genome-wide SNP genotyping data, six Klinefelter syndrome (47XXY) subjects were identified (Additional file 1: Figure S4). Five of the six Klinefelter subjects met criteria for PRISm by fixed threshold criteria while 3 met criteria for PRISm by LLN standards. This represents a significant enrichment of PRISm among Klinefelter syndrome subjects regardless of whether fixed threshold or LLN criteria were applied (Fisher's exact p-values 1.53×10^{-4} and 0.02, respectively).

We performed a genome-wide association study (GWAS) of PRISm status relative to control subjects; results for the analysis in fixed threshold-defined cohorts are illustrated in Additional file 1: Figure S5. Although no genetic variant met the genome-wide threshold for significance (p-value < 5×10^{-8}) in either the fixed threshold (Additional file 1: Table S4) or LLN – defined analyses (Additional file 1: Table S5), several SNPs with suggestive p-values were identified within the pleckstrin homology domain containing, family A member 5 (*PLEKHA5*) gene as well as within the voltage-dependent L-type calcium channel subunit beta-2 (*CACNB2*) gene. The most highly associated SNP from the fixed-threshold analysis (rs113840005 in *PLEKHA5*) was among the top 10 variants identified in the LLN-analysis. Considerable heterogeneity, as illustrated in the 10th – 90th percentile values for selected variables (Additional file 1: Table S6), exists among PRISm subjects and may contribute to the lack of genetic signal in GWAS analysis.

Aim 3: Unsupervised cluster analysis to identify potential subgroups within PRISm subjects

The subset of fixed threshold-defined PRISm subjects with complete data included in the k-means clustering analysis (n = 1,135) did not differ from the full cohort of PRISm subjects with respect to mean age, pack-years, BMI, or distribution by gender or current smoking status. Normalized mutual information (NMI) analysis using a five-fold cross-validation strategy demonstrated high cluster reproducibility for k = 3 clusters (Additional file 1:

Table S7). An overview of the analysis is shown qualitatively in Figure 2 while the specific results of unsupervised k-means clustering in non-Hispanic white and African American PRISm subjects are illustrated in Figures 3 and 4, respectively. Clusters observed in NHW were reasonably reproducible in the AA, as illustrated in Additional file 1: Figure S6. Subgroup characteristics by cluster are summarized in Table 3. Members of Cluster 1 demonstrate the highest FEV₁/FVC ratio and forced expiratory flow rate at 25%-75% of FVC (FEF₂₅₋₇₅), as well as the lowest mean emphysema and gas trapping; we refer to this cluster as a putative “PRISm – Restricted cluster”. Members of Cluster 2 have the lowest mean BMI and FEV₁/FVC ratio, as well as the highest mean emphysema and gas trapping; we refer to this cluster as the “PRISm – COPD cluster”. Finally, members of Cluster 3 have the highest mean BMI, the greatest degree of impairment in FEV₁% predicted, the thickest segmental wall area, the lowest FEF₂₅₋₇₅ flow rates, and the highest prevalence of diabetes mellitus and rates of bronchodilator responsiveness (BDR); we refer to this cluster as the “PRISm – Metabolic cluster”. Members of the Cluster 3 also demonstrate the highest mean MMRC dyspnea scores and the lowest mean 6 minute walk distance.

We also performed cluster analysis on the LLN-defined PRISm cohort; the subset of subjects with complete data (n = 978) did not differ from the full LLN-PRISm cohort with respect to mean age, pack-years smoked, BMI, or distribution by gender; there were significantly fewer current smokers (60% vs. 70.2%) in the subset with complete data used for cluster analysis. NMI and silhouette width analysis demonstrated high cluster reproducibility for k = 4 clusters (Additional file 1: Table S7). The results of unsupervised k-means clustering in NHW and AA subjects are illustrated in Additional file 1: Figure S7 (panels (a) and (b) respectively). Separation between clusters in each of these analyses (NHW and AA) is not as distinct as in the fixed threshold analysis; additionally, the clusters found in NHW did not appear to overlap well with clusters identified in the AA analysis (Additional file 1: Figure S8). Characteristics of each cluster are summarized in Additional file 1: Table S9. We have putatively assigned Clusters 1, 2, and 3 to be analogous to the “PRISm-restrictive”, “PRISm-COPD”, and “PRISm-metabolic” subtypes described in the fixed-threshold analysis. Members of Cluster 4 have the highest rates of current smoking; however, beyond that, the clusters appear to represent relatively distinct subgroups in NHW and AA subjects. In the NHW LLN-PRISm Cluster 4, subjects have the lowest BMI and highest resting oxygen saturation and best exercise capacity while Cluster 4 subjects in the AA LLN-PRISm analysis appear to be more ill with the greatest impairment in FEV₁% predicted, increased segmental wall area thickness and decreased FEF₂₅₋₇₅.

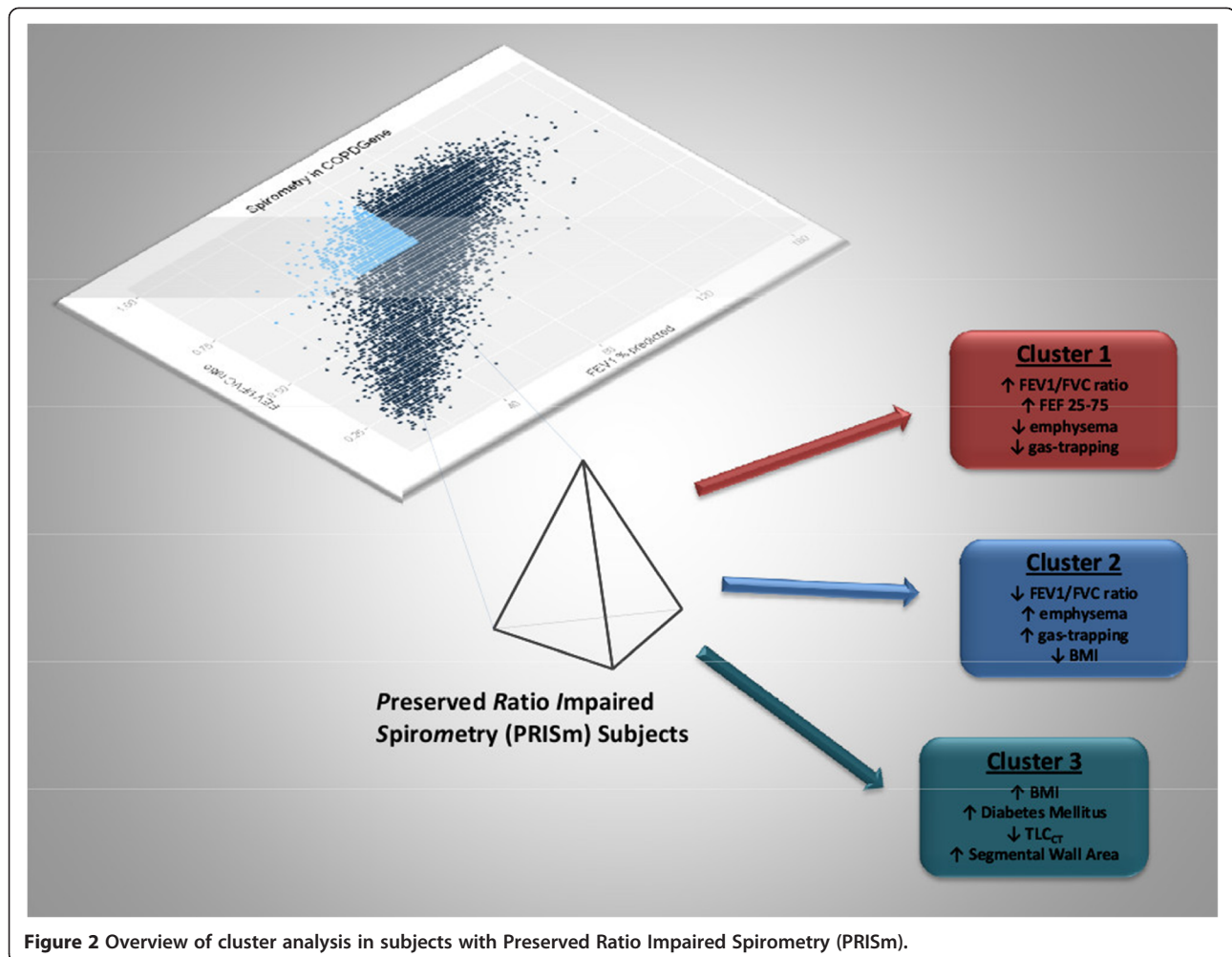


Figure 2 Overview of cluster analysis in subjects with Preserved Ratio Impaired Spirometry (PRISm).

Genetic variants previously described in studies of COPD, interstitial lung disease, and metabolic phenotypes were examined for associations with the clusters identified in the fixed-threshold analysis (Additional file 1: Table S10). The minor (risk) allele frequency of rs8050136, located in the first intron of the fat mass and obesity associated (*FTO*) gene, by subgroup is illustrated in Additional file 1: Figure S9; a relative enrichment of risk alleles in the PRISm – Metabolic subgroup was noted among African American subjects (ANOVA p-value 0.05), however, this enrichment was not statistically significant among non-Hispanic white subjects.

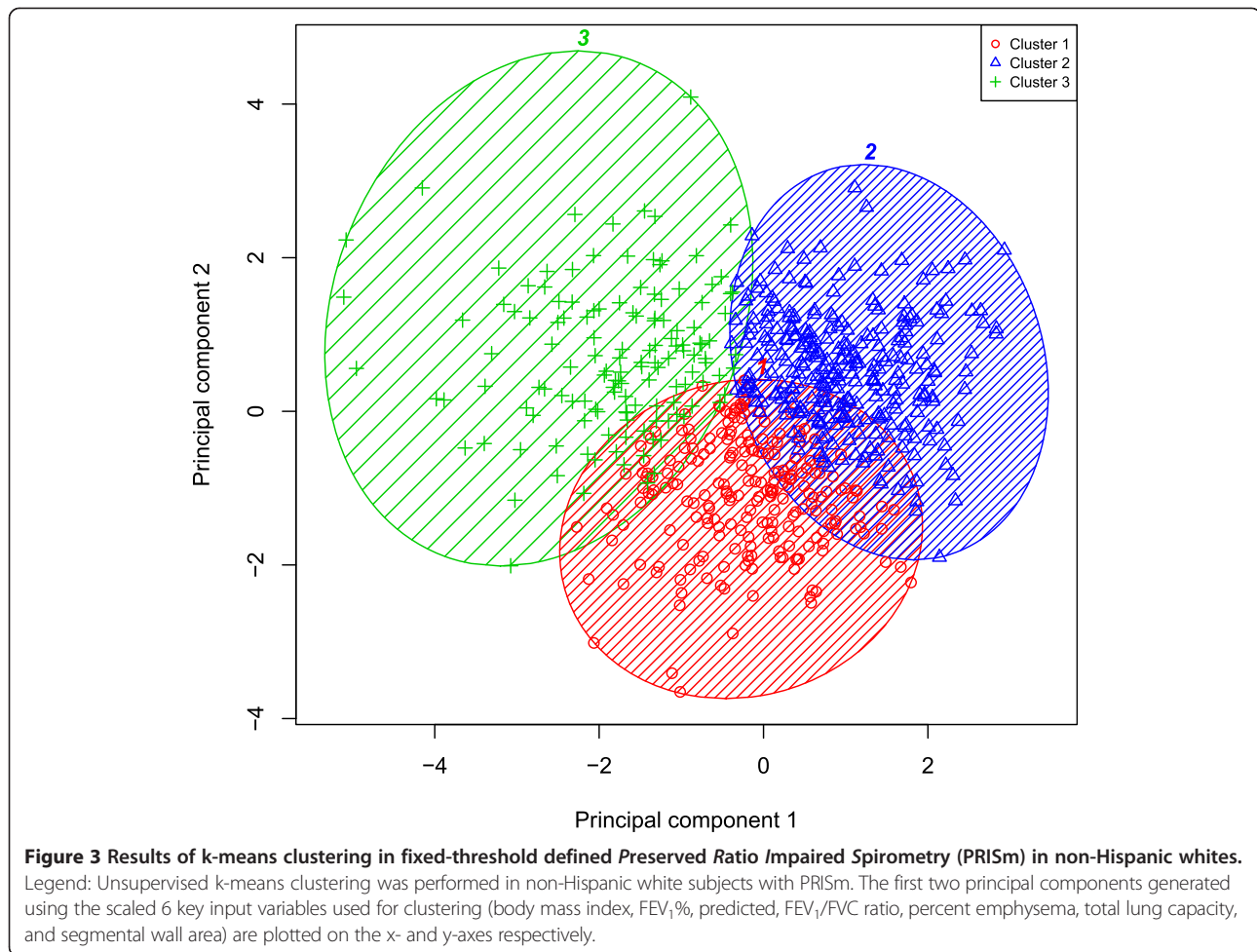
Discussion

In this manuscript, we examine detailed demographic, spirometric, and radiographic features of subjects with Preserved Ratio Impaired Spirometry and leverage these data to explore genetic associations and subgroups within the cohort. We confirm the overall prevalence of PRISm within current and former smokers in our cohort

is consistent with the prevalence reported in other cross-sectional studies, including several population-based studies [10–13,15,31]. We affirm previously reported associations with body mass index and diabetes mellitus and report novel associations with radiographic and functional predictors of PRISm status (Aim 1). While no genome-wide significant genetic predictors were identified in our GWAS studies, we uncovered a novel association between PRISm and Klinefelter's syndrome (Aim 2). Finally, the results of unsupervised clustering analysis demonstrate 3 clusters which may represent pathobiologically distinct subgroups within the PRISm cohort.

Aim 1: Epidemiology of PRISm

As with obstructive lung diseases, controversy regarding the delineation between normal and abnormal exists for PRISm. Differences in the populations defined by fixed threshold (i.e. GOLD criteria) versus lower limit of normal FEV₁ criteria likely contribute to differences in associations with certain epidemiological predictors identified in



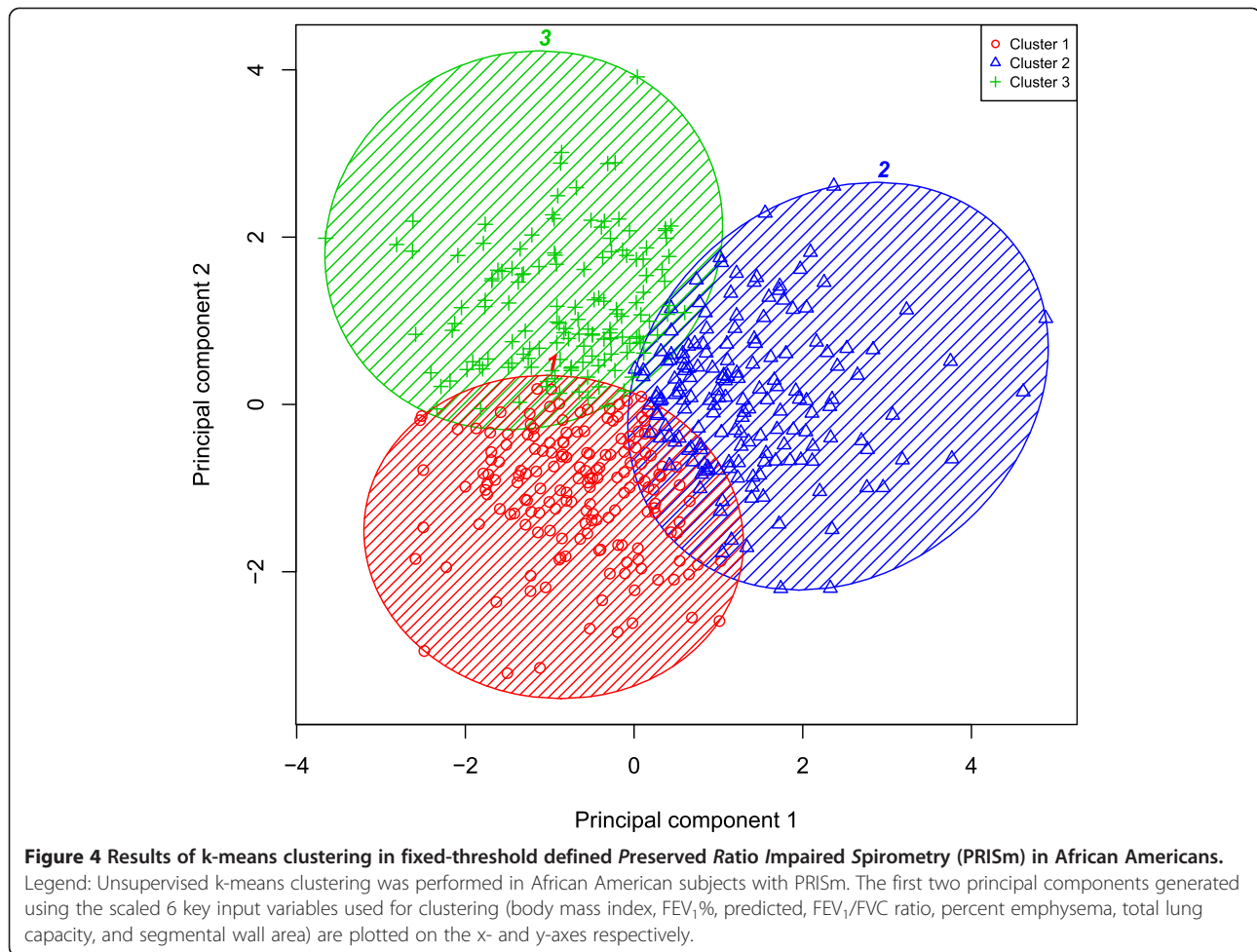
our analysis; for example, the enrichment of African Americans in the LLN-defined PRISm cohort may reflect less accurate population-based prediction equations for or increased variability in African Americans rather than a distinct pathobiological process.

Despite the lack of a consensus definition for PRISm, the majority of associations reported in our study were remarkably consistent regardless of whether fixed (GOLD) or LLN criteria were utilized. Previously reported associations with increased mean BMI and a high prevalence of comorbid conditions such as diabetes mellitus [6,9,10,12,13,15,18] were observed in our cohort on univariate analyses. We additionally confirm associations with decreased total lung capacity and decreased emphysema first reported in our analysis of PRISm subjects among the first 2500 subjects recruited in COPDGene [8].

In multivariable models, PRISm subjects had increased cumulative exposure to tobacco smoke as well as an increased prevalence of physician-diagnosed asthma and peripheral vascular disease relative to control subjects. These factors may contribute to the increase in symptoms

as assessed through the MMRC score, decreased exercise tolerance, and decreased resting oxygen saturation also observed in this cohort relative to control subjects. In multivariate models comparing PRISm with COPD subjects, increased body mass index and an increased prevalence of diabetes mellitus were consistently identified as predictors; whether these factors are pathobiologically related to the development of these two distinct disease states is a topic that warrants investigation in the future. Radiographic variables, such as percent emphysema and TLC, were among the most consistently identified predictors of PRISm status in multivariate models relative to *both* control and COPD subjects.

The role of increased BMI among PRISm subjects continues to deserve special consideration. Although obesity has been associated with proportionate decreases in FEV₁ and FVC as well as decreases in TLC, lung function values of obese subjects typically remain within the normal range [32]; thus the degree of impairment in lung function in PRISm subjects is unlikely to be due solely to the mechanical properties of increased body mass. This supposition is supported by the divergent



trends in the prevalence of overweight and obesity relative to PRISm over the last half century (Additional file 1: Figure S10); the dramatic increase in the prevalence of overweight and obesity [33] is not reflected in the relatively stable prevalence of PRISm [7–9,11–13,15,16,22,31,34].

Aim 2: Genetic associations

The association between Klinefelter's syndrome and PRISm is consistent with previous reports of an increased prevalence of restriction [35–37] in this population; in fact, all six Klinefelter subjects identified demonstrated a TLC_{CT} < 80% predicted (data not shown). We acknowledge the limitations associated with the use of male prediction equations for lung function in this subgroup, as traditional formulas do not account for the eunuchoid proportions which characterize this syndrome. However, given previous reports of decreased lung compliance [36], an increased prevalence of respiratory symptoms [38,39], and increased mortality due to respiratory causes among Klinefelter's subjects [40,41], we believe the association may be indicative of true pathobiological differences and warrants additional investigation in the future.

Although no single genetic variant met the accepted genome-wide threshold for significance, we identified suggestive associations between PRISm and variants within the *PLEKHA5* and *CACNB2* genes. rs113840005, which was a top variant in both the fixed threshold and LLN analyses, is located within an intron of the *PLEKHA5* gene. Multiple splice variants of this gene have been identified; some isoforms are ubiquitously expressed while other isoforms are highly specific to tissue type and developmental stage. In general, the protein products of the *PLEKHA5* gene are typically located in the cytosol of cells and are believed to contribute to intracellular signaling and cytoskeletal organization [42]; the mechanism by which variants in this gene are associated with PRISm is unclear. Intronic variants within the *CACNB2* gene were among the most strongly associated in the fixed-threshold analysis. Analogous to *PLEKHA5*, multiple isoforms of the protein product exist. Variants within this gene have been associated with blood pressure levels and hypertension [43,44] as well as Brugada syndrome [45–47]. Additional investigations into the mechanism behind the association of variants at this locus with PRISm are warranted.

Table 3 Results of unsupervised k-means clustering in (a) non-Hispanic white and (b) African American subjects with fixed threshold-defined Preserved Ratio Impaired Spirometry (PRISm)

| Feature | (a) Non-Hispanic White subjects | | | (b) African American subjects | | |
|--------------------------------------|---------------------------------|----------------|----------------|-------------------------------|----------------|----------------|
| | Cluster 1 | Cluster 2 | Cluster 3 | Cluster 1 | Cluster 2 | Cluster 3 |
| n | 227 | 291 | 137 | 171 | 167 | 142 |
| Age | 57.9 (8.2) | 62.2 (8.1) | 59.2 (9.4) | 52.9 (5.8) | 54.6 (6.1) | 52.6 (5.3) |
| Sex (% male) | 34.8 | 45.7 | 51.1 | 47.4 | 56.9 | 45.1 |
| Body Mass Index | 30.9 (6.4) | 30.4 (6.3) | 36.2 (7.1) | 32.0 (6.8) | 27.3 (6.0) | 35.7 (8.4) |
| Current smoker | 53.3 | 43.6 | 49.6 | 84.8 | 83.8 | 80.3 |
| Pack-years | 42.8 (22.6) | 46.6 (24.5) | 50.9 (29.7) | 35.6 (18.5) | 39.6 (24.0) | 38.1 (21.4) |
| FEV ₁ % predicted | 74.2 (4.3) | 73.4 (5.3) | 60.5 (7.8) | 73.3 (5.1) | 72.7 (6.2) | 62.5 (8.9) |
| FEV ₁ /FVC | 0.80 (0.04) | 0.74 (0.03) | 0.74 (0.04) | 0.83 (0.04) | 0.74 (0.03) | 0.75 (0.03) |
| FEF ₂₅₋₇₅ | 2.21 (0.65) | 1.58 (0.47) | 1.41 (0.47) | 2.38 (0.71) | 1.58 (0.50) | 1.42 (0.43) |
| Resting O ₂ saturation | 96.4(2.3) | 96.3 (2.1) | 95.2 (2.9) | 97.1 (2.2) | 97.1 (2.3) | 96.6 (2.7) |
| TLC _{CT} *% predicted | 79.4 (9.8) | 90.4 (10.7) | 74.4 (12.7) | 71.0 (11.1) | 83.6 (12.3) | 71.6 (12.3) |
| Percent emphysema | 0.6 (0.7) | 2.6 (2.7) | 0.9 (1.1) | 0.5 (0.6) | 2.7 (4.5) | 0.8 (1.0) |
| Segmental Wall Area% | 62.0 (2.5) | 60.6 (2.3) | 64.7 (2.6) | 62.4 (2.8) | 61.9 (2.9) | 65.7 (2.5) |
| Percent Gas Trapping | 6.5 (6.0) | 13.5 (8.6) | 9.8 (7.1) | 7.2 (7.3) | 14.8 (13.5) | 9.4 (7.5) |
| Diabetes Mellitus | 18.5 | 14.4 | 36.5 | 22.8 | 15.6 | 27.5 |
| Hypertension | 36.6 | 47.4 | 56.2 | 53.8 | 51.5 | 54.2 |
| Hyperlipidemia | 45.4 | 58.8 | 59.9 | 31.0 | 24.0 | 29.6 |
| Chronic bronchitis | 16.3 | 15.8 | 32.9 | 11.1 | 19.2 | 14.8 |
| Bronchodilator Response [†] | 11.5 | 13.5 | 22.2 | 11.4 | 9.2 | 20.1 |
| MMRC [‡] | 1.154 (1.32) | 1.13 (1.29) | 1.95 (1.50) | 1.63 (1.53) | 1.57 (1.47) | 1.96 (1.53) |
| 6 minute walk distance | 1386.5 (347.1) | 1392.2 (320.4) | 1132.6 (390.2) | 1219.9 (349.4) | 1279.5 (338.1) | 1099.2 (342.1) |

Results are reported as mean (SD) or percent.

Preserved Ratio Impaired Spirometry (PRISm) defined as: FEV₁/FVC ≥ 0.7 & FEV₁ < 80% predicted.

*TLC_{CT} = Total Lung Capacity by computed tomography.

[†]Bronchodilator response considered present if the change in FEV₁ or FVC was >200 mL and ≥ 12% predicted following administration of short acting inhaled beta-agonist.

[‡]MMRC = Modified Medical Research Council.

Aim 3: Subgroup identification

Clinical management strategies for PRISm are poorly defined and reflect the low specificity of spirometric measurements alone in identifying distinct disease processes in this cohort [13,14]. Overt and subclinical interstitial lung disease [5], chest wall abnormalities, neuromuscular and functional impairments, as well as airway diseases (such as asthma and chronic obstructive pulmonary disease) which have classically been associated with obstruction [6,9,48], can all produce the PRISm pattern on spirometry; what remains unknown is the proportion attributable to each process and how to identify different groups of subjects. When we incorporated key clinical and radiographic variables with existing spirometric data into an unbiased clustering algorithm, we were able to identify clinically relevant subtypes within the PRISm cohort.

Previous studies have supported the existence of an airway disease/COPD subgroup among PRISm subjects [6,9]; we were likewise able to identify a putative COPD

subgroup (Cluster 2 in both the NHW and AA analyses) with evidence of relatively increased emphysema and gas trapping, preserved TLC, and a relatively reduced FEV₁/FVC ratio. Interestingly, this group appears to have the lowest degree of physiological impairment; they experience the least dyspnea (as assessed by MMRC score), have the best exercise capacity (highest 6MWD), and the least hypoxemia. In a longitudinal study by Guerra et al. [9], approximately one-third of subjects with PRISm eventually developed airflow obstruction on spirometry – Cluster 2 may be enriched for subjects with “early COPD” who have not yet developed the classical obstructive pattern. Longitudinal data, which is currently being collected in the COPDGene cohort, will be crucial to investigating this hypothesis.

The PRISm-metabolic subgroup represents a highly symptomatic and functionally limited group for whom treatment options have not been systematically explored. This subgroup has the greatest degree of spirometric impairment in FEV₁ which may be related to the increased

subsegmental airway wall thickness. Given these findings, as well as the high prevalence of bronchodilator responsiveness in this subgroup, the benefit of inhaled steroids and/or bronchodilators in this subgroup presents a clinically relevant question for future studies.

In summary, we have analyzed the epidemiological and radiographic predictors, explored clinically relevant putative subgroups, and identified a novel association with Klinefelter's syndrome in PRISm. The strengths of the current study include the utilization of a large cohort with rich data in multiple domains as well as the application of rigorous, unbiased interrogations to both characterize and subtype this relatively understudied syndrome. Despite this, we acknowledge the following limitations. First, the lack of visual assessments of CT data for the majority of the cohort limits our ability to ascertain the impact of chest wall or diaphragmatic abnormalities and atypical interstitial/parenchymal infiltrates among the PRISm cohort. Second, although this cohort is the largest to date with genetic data available, the number of subjects is modest in the context of genome-wide association studies and limits our power to detect variants of modest effect sizes. Lastly, the degree to which the findings reported in our study are generalizable to other populations, such as non-smokers and subjects outside of the United States, should be explored. Future work in independent populations of PRISm subjects, as well as *in vivo* and *in vitro* work in model systems, to explore the biological mechanisms behind the associations reported in our manuscript are warranted.

Additional file

Additional file 1: Additional Methods - Additional information on cohorts, variables, and analyses. **Table S1** - Respiratory medication use by physician diagnosed asthma and BDR, **Table S2**-Characteristics of LLN-defined cohorts, **Table S3** - Multivariate predictors of LLN-PRISm, **Table S4**-Top 10 associations of Fixed-threshold PRISm vs. Controls GWAS, **Table S5**-Top 10 associations of LLN-PRISm vs. Controls GWAS, **Table S6**- 10th-90th percentiles of variables in PRISm, **Table S7**-Normalized mutual information and silhouette width for fixed-threshold PRISm cluster analysis, **Table S8** - Normalized mutual information and silhouette width for LLN-PRISm cluster analysis, **Table S9**-Results of unsupervised k-means clustering in LLN-PRISm, **Table S10**-Candidate gene testing in fixed-threshold PRISm subgroups. **Figure S1** - Distribution of spirometry by LLN criteria, **Figure S2** - Overlap between fixed-threshold and LLN PRISm, **Figure S3** - Reclassification of PRISm subjects by fixed-threshold and LLN criteria, **Figure S4** - Identification of Klinefelter subjects by X and Y chromosome intensity analysis, **Figure S5** - QQ and Manhattan plots of fixed-threshold PRISm GWAS, **Figure S6** - Superimposed K-means clustering results in NHW+AA subjects (fixed-threshold PRISm), **Figure S7** - K-means clustering results in LLN-PRISm, **Figure S8** - Superimposed K-means clustering results in NHW + AA, **Figure S9** - MAF of rs8050136 by fixed-threshold PRISm subgroup, **Figure S10** - Prevalence of overweight/obesity and PRISm (1960-2010).

Abbreviations

PRISm: Preserved ratio impaired spirometry; FEF₂₅₋₇₅: Forced expiratory flow rate at 25%-75% of forced vital capacity; FEV1: Forced expiratory volume in the first second; FVC: Forced vital capacity; GWAS: Genome wide association

study; GOLD: Global initiative for obstructive lung disease; %LAA-950_{insp}: Percent low attenuation areas at -950 Hounsfield units on inspiratory CT scan; LLN: Lower limits of normal; BMI: Body mass index; NHW: Non-Hispanic white; NMI: Normalized mutual information; AA: African American; CT: Computed tomography; COPD: Chronic obstructive pulmonary disease; SNP: Single nucleotide polymorphism; BDR: Bronchodilator responsiveness; MMRC: Modified Medical Research Council; TLC: Total lung capacity.

Competing interests

Emily S. Wan, Douglas Curran-Everett, Elizabeth A. Regan, Jeffrey L. Curtis, Dawn L. DeMeo, Peter J. Castaldi, Terri H. Beaty, John E. Hokanson, and James D. Crapo have no conflicts to disclose.

Edwin K. Silverman received grant support from GlaxoSmithKline for studies of COPD genetics and has received honoraria and consulting fees from AstraZeneca, Merck, and GlaxoSmithKline.

Michael H. Cho has received consulting fees from Merck.

MeiLan K. Han has participated in advisory boards for Boehringer Ingelheim GmbH, Pfizer, GlaxoSmithKline, Novartis, Forest and MedImmune, has participated on speaker's bureaus for Boehringer Ingelheim GmbH, Pfizer, GlaxoSmithKline, Forest, Grifols, the National Association for Continuing Education, and WebMD, has served as a consultant for Novartis, Nycomed and Ikaria and has received royalties from UpToDate and ePocrates, Inc. David A. Lynch's institution and laboratory receives research support from the National Heart Lung and Blood Institute, Siemens, Inc, Perceptive Imaging, Inc, and Centocor, Inc, Inc. Dr Lynch is a consultant to Perceptive Imaging, Inc, Boehringer Ingelheim, Inc, Genentech, Inc, Gilead, Inc, and Intermune, Inc. Related to the general topic area of COPD over the last 3 years, Barry J. Make has served on medical advisory boards for AstraZeneca, Sunovion, Boehringer-Ingelheim, GlaxoSmithKline, Forest, Ikaria, Aerocrine, Novartis, MedImmune, Covidien, Breathe. Funds for multi-center studies have been provided to and controlled by National Jewish Health from NHLBI, AstraZeneca, Sunovion, Boehringer-Ingelheim, GlaxoSmithKline, Forest, NABI. He has been on the Speaker's Bureaus for GlaxoSmithKline, Boehringer-Ingelheim, GlaxoSmithKline and Forest.

Authors' contributions

All authors participated in drafting and revising the submitted manuscript. E. S.W., P.J.C., M.H.C., and T.H.B. contributed to the analysis and interpretation of data. J.E.H., B.J.M., D.A.L., D.L.D., J.D.C., and E.K.S. contributed to the conception and design of the study. E.A.R., M.K.H., J.L.C., and D.C.E. contributed to the acquisition of data. All authors read and approved the final manuscript.

Acknowledgements

We acknowledge and thank the COPDGene Core Teams:

Administrative Core: James Crapo, MD (PI), Edwin Silverman, MD, PhD (PI), Barry Make, MD, Elizabeth Regan, MD, PhD, Rochelle Lantz, Lori Stepp, Sandra Melanson,

Genetic Analysis Core: Terri Beaty, PhD, Barbara Klanderman, PhD, Nan Laird, PhD, Christoph Lange, PhD, Michael Cho, MD, Stephanie Santorico, PhD, John Hokanson, MPH, PhD, Dawn DeMeo, MD, MPH, Nadia Hansel, MD, MPH, Craig Hersh, MD, MPH, Peter Castaldi, MD, MSc, Merry-Lynn McDonald, PhD, Jin Zhou, MD, PhD, Manuel Mattheissen, MD, PhD, Emily Wan, MD, Megan Hardin, MD, Jacqueline Hetmanski, MS, Margaret Parker, MS, Tanda Murray, MS
Imaging Core: David Lynch, MB, Joyce Schroeder, MD, John Newell, Jr., MD, John Reilly, MD, Harvey Coxson, PhD, Philip Judy, PhD, Eric Hoffman, PhD, George Washko, MD, Raul San Jose Estepar, PhD, James Ross, MSc, Mustafa Al Qaisi, MD, Jordan Zach, Alex Kluiber, Jered Sieren, Tanya Mann, Deanna Richert, Alexander McKenzie, Jaleh Akhavan, Douglas Stinson
PFT QA Core, LDS Hospital, Salt Lake City, UT: Robert Jensen, PhD
Biological Repository, Johns Hopkins University, Baltimore, MD: Homayoon Farzadegan, PhD, Stacey Meyerer, Shivam Chandan, Samantha Bragan
Data Coordinating Center and Biostatistics, National Jewish Health, Denver, CO: Douglas Everett, PhD, Andre Williams, PhD, Carla Wilson, MS, Anna Forssen, MS, Amber Powell, Joe Piccoli

Epidemiology Core, University of Colorado School of Public Health, Denver, CO: John Hokanson, MPH, PhD, Marci Sontag, PhD, Jennifer Black-Shinn, MPH, Gregory Kinney, MPH, PhDc, Sharon Lutz, MPH, PhD.

We further wish to acknowledge the COPDGene Investigators from the participating Clinical Centers:

Ann Arbor VA: Jeffrey Curtis, MD, Ella Kazerooni, MD

Baylor College of Medicine, Houston, TX: Nicola Hanania, MD, MS, Philip Alapat, MD, Venkata Bandi, MD, Kalpalatha Guntupalli, MD, Elizabeth Guy, MD, Antara Mallampalli, MD, Charles Trinh, MD, Mustafa Atik, MD, Hasan Al-Azzawi, MD, Marc Willis, DO, Susan Pinero, MD, Linda Fahr, MD, Arun Nachiappan, MD, Collin Bray, MD, L. Alexander Frigini, MD, Carlos Farinas, MD, David Katz, MD, Jose Freytes, MD, Anne Marie Marciel, MD
Brigham and Women's Hospital, Boston, MA: Dawn DeMeo, MD, MPH, Craig Hersh, MD, MPH, George Washko, MD, Francine Jacobson, MD, MPH, Hiroto Hatabu, MD, PhD, Peter Clarke, MD, Ritu Gill, MD, Andetta Hunsaker, MD, Beatrice Trotman-Dickenson, MBBS, Rachna Madan, MD
Columbia University, New York, NY: R. Graham Barr, MD, DrPH, Byron Thomashow, MD, John Austin, MD, Belinda D'Souza, MD
Duke University Medical Center, Durham, NC: Neil MacIntyre, Jr., MD, Lacey Washington, MD, H Page McAdams, MD
Fallon Clinic, Worcester, MA: Richard Rosiello, MD, Timothy Bresnahan, MD, Joseph Bradley, MD, Sharon Kuong, MD, Steven Meller, MD, Suzanne Roland, MD
Health Partners Research Foundation, Minneapolis, MN: Charlene McEvoy, MD, MPH, Joseph Tashjian, MD
Johns Hopkins University, Baltimore, MD: Robert Wise, MD, Nadia Hansel, MD, MPH, Robert Brown, MD, Gregory Diette, MD, Karen Horton, MD
Los Angeles Biomedical Research Institute at Harbor UCLA Medical Center, Los Angeles, CA: Richard Casaburi, MD, Janos Porszasz, MD, PhD, Hans Fischer, MD, PhD, Matt Budoff, MD, Mehdi Rambod, MD
Michael E. DeBakey VAMC, Houston, TX: Amir Sharafkhaneh, MD, Charles Trinh, MD, Hirani Kamal, MD, Roham Darvishi, MD, Marc Willis, DO, Susan Pinero, MD, Linda Fahr, MD, Arun Nachiappan, MD, Collin Bray, MD, L. Alexander Frigini, MD, Carlos Farinas, MD, David Katz, MD, Jose Freytes, MD, Anne Marie Marciel, MD
Minneapolis VA: Dennis Niewoehner, MD, Quentin Anderson, MD, Kathryn Rice, MD, Audrey Caine, MD
Morehouse School of Medicine, Atlanta, GA: Marilyn Foreman, MD, MS, Gloria Westney, MD, MS, Eugene Berkowitz, MD, PhD
National Jewish Health, Denver, CO: Russell Bowler, MD, PhD, David Lynch, MB, Joyce Schroeder, MD, Valerie Hale, MD, John Armstrong, II, MD, Debra Dyer, MD, Jonathan Chung, MD, Christian Cox, MD
Temple University, Philadelphia, PA: Gerard Criner, MD, Victor Kim, MD, Nathaniel Marchetti, DO, Aditi Satti, MD, A. James Mamary, MD, Robert Steiner, MD, Chandra Dass, MD, Libby Cone, MD
University of Alabama, Birmingham, AL: William Bailey, MD, Mark Dransfield, MD, Michael Wells, MD, Surya Bhatt, MD, Hrudaya Nath, MD, Satinder Singh, MD
University of California, San Diego, CA: Joe Ramsdell, MD, Paul Friedman, MD
University of Iowa, Iowa City, IA: Alejandro Cornellias, MD, John Newell, Jr., MD, Edwin JR van Beek, MD, PhD
University of Michigan, Ann Arbor, MI: Fernando Martinez, MD, MeiLan Han, MD, Ella Kazerooni, MD
University of Minnesota, Minneapolis, MN: Christine Wendt, MD, Tadashi Allen, MD
University of Pittsburgh, Pittsburgh, PA: Frank Sciruba, MD, Joel Weissfeld, MD, MPH, Carl Fuhrman, MD, Jessica Bon, MD, Danielle Hooper, MD
University of Texas Health Science Center at San Antonio, San Antonio, TX: Antonio Anzueto, MD, Sandra Adams, MD, Carlos Orozco, MD, Mario Ruiz, MD, Amy Mumbower, MD, Ariel Kruger, MD, Carlos Restrepo, MD, Michael Lane, MD
We additionally thank Margaret Parker for her assistance in generating figures for the manuscript.

Funding/Support and Roles of Sponsors

NIH R01 HL089856 and NIH P01 HL105339 (E.K.S.), NIH R01 HL089897 (J.D.C.), NIH K12 089990 (E.S.W.), Brigham & Women's Center for Faculty Development and Diversity Career Development Award (E.S.W.), and a Parker B. Francis Foundation Fellowship (E.S.W.) The COPDGene® project is also supported by the COPD Foundation through contributions made to an Industry Advisory Board comprised of AstraZeneca, Boehringer Ingelheim, Novartis, Pfizer, Siemens, GlaxoSmithKline, and Sunovion.

The content of this manuscript is solely the responsibility of the authors; none of the above named entities participated in the design or conduct of the study; the collection, management, analysis, or interpretation of the data; or the preparation, review, approval, or decision to submit the manuscript for publication.

Author details

¹Channing Division of Network Medicine, Brigham and Women's Hospital, 181 Longwood Avenue, 4th floor, Boston, MA 2115, USA. ²Division of

Pulmonary and Critical Care Medicine, Brigham and Women's Hospital, Boston, MA, USA. ³Department of Epidemiology, Colorado School of Public Health, University of Colorado, Denver, CO, USA. ⁴National Jewish Health, Denver, CO, USA. ⁵Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA. ⁶Division of Pulmonary and Critical Care Medicine, University of Michigan Health System, Ann Arbor, MI, USA. ⁷Pulmonary and Critical Care Section, Ann Arbor Veterans Affairs Healthcare System, Ann Arbor, MI, USA. ⁸Division of Biostatistics and Bioinformatics, National Jewish Health, Denver, CO, USA. ⁹Department of Biostatistics and Informatics, Colorado School of Public Health, University of Colorado Denver, Denver, CO, USA.

Received: 20 February 2014 Accepted: 27 July 2014

Published: 6 August 2014

References

1. Petty TL: John Hutchinson's mysterious machine revisited. *Chest* 2002, **121** (5 Suppl):219S–223S.
2. Aaron SD, Dales RE, Cardinal P: How accurate is spirometry at predicting restrictive pulmonary impairment? *Chest* 1999, **115**(3):869–873.
3. Swanney MP, Beckert LE, Frampton CM, Wallace LA, Jensen RL, Crapo RO: Validity of the American Thoracic Society and other spirometric algorithms using FVC and forced expiratory volume at 6 s for predicting a reduced total lung capacity. *Chest* 2004, **126**(6):1861–1866.
4. Vandevoorde J, Verbanck S, Schuermans D, Broekaert L, Devroey D, Kartounian J, Vincken W: Forced vital capacity and forced expiratory volume in six seconds as predictors of reduced total lung capacity. *Eur Respir J* 2008, **31**(2):391–395.
5. Washko GR, Hunninghake GM, Fernandez IE, Nishino M, Okajima Y, Yamashiro T, Ross JC, Estepar RS, Lynch DA, Brehm JM, Adnriole KP, Diaz AA, Khorasani R, D'Aco K, Sciruba FC, Silverman EK, Hatabu H, Rosas IO, COPDGene Investigators: Lung volumes and emphysema in smokers with interstitial lung abnormalities. *N Engl J Med* 2011, **364**(10):897–906.
6. Iyer VN, Schroeder DR, Parker KO, Hyatt RE, Scanlon PD: The nonspecific pulmonary function test: longitudinal follow-up and outcomes. *Chest* 2010, **139**(4):878–886.
7. Kohler D, Fischer J, Raschke F, Schonhofer B: Usefulness of GOLD classification of COPD severity. *Thorax* 2003, **58**(9):825.
8. Wan ES, Hokanson JE, Murphy JR, Regan EA, Make BJ, Lynch DA, Crapo JD, Silverman EK: Clinical and radiographic predictors of GOLD-unclassified smokers in the COPDGene study. *Am J Respir Crit Care Med* 2011, **184**(1):57–63.
9. Guerra S, Sherrill DL, Venker C, Ceccato CM, Halonen M, Martinez FD: Morbidity and mortality associated with the restrictive spirometric pattern: a longitudinal study. *Thorax* 2010, **65**(6):499–504.
10. Mannino DM, Buist AS, Petty TL, Enright PL, Redd SC: Lung function and mortality in the United States: data from the First National Health and Nutrition Examination Survey follow up study. *Thorax* 2003, **58**(5):388–393.
11. Mannino DM, Doherty DE, Sonia Buist A: Global Initiative on Obstructive Lung Disease (GOLD) classification of lung disease and mortality: findings from the Atherosclerosis Risk in Communities (ARIC) study. *Respir Med* 2006, **100**(1):115–122.
12. Mannino DM, Ford ES, Redd SC: Obstructive and restrictive lung disease and functional limitation: data from the Third National Health and Nutrition Examination. *J Intern Med* 2003, **254**(6):540–547.
13. Mannino DM, McBurnie MA, Tan W, Kocabaş A, Anto J, Vollmer WM, Buist AS: Restricted spirometry in the burden of lung disease study. *Int J Tuberc Lung Dis* 2012, **16**(10):1405–1411.
14. Quanjer PH, Brazzale DJ, Boros PW, Pretto JJ: Implications of adopting the Global Lungs Initiative 2012 all-age reference equations for spirometry. *Eur Respir J* 2013, **42**(4):1046–1054.
15. Soriano JB, Miravittles M, Garcia-Rio F, Munoz L, Sanchez G, Sobradillo V, Duran E, Guerrero D, Ancochea J: Spirometrically-defined restrictive ventilatory defect: population variability and individual determinants. *Prim Care Respir J* 2012, **21**(2):187–193.
16. Vaz Fragoso CA, Gill TM, McAvay G, Yaggy HK, Van Ness PH, Concato J: Respiratory impairment and mortality in older persons: a novel spirometric approach. *J Invest Med* 2011, **59**(7):1089–1095.
17. Global strategy for the diagnosis, management and prevention of COPD. <http://www.goldcopd.org>.

18. Mannino DM, Holguin F, Pavlin BI, Ferdinands JM: **Risk factors for prevalence of and mortality related to restriction on spirometry: findings from the First National Health and Nutrition Examination Survey and follow-up.** *Int J Tuberc Lung Dis* 2005, **9**(6):613–621.
19. Wan ES, Cho MH, Boutaoui N, Klanderman BJ, Sylvia JS, Ziniti JP, Won S, Lange C, Pillai SG, Anderson WH, Kong X, Lomas DA, Bakke PS, Gulsvik A, Regan EA, Murphy JR, Make BJ, Crapo JD, Wouters EF, Celli BR, Silverman EK, DeMeo DL, ECLIPSE investigators, Norway-Bergen cohort, National Emphysema Treatment Trial, and the COPDGene Investigators: **Genome-wide association analysis of body mass in chronic obstructive pulmonary disease.** *Am J Respir Cell Mol Biol* 2011, **45**(2):304–310.
20. Mannino DM, Ford ES, Redd SC: **Obstructive and restrictive lung disease and markers of inflammation: data from the Third National Health and Nutrition Examination.** *Am J Med* 2003, **114**(9):758–762.
21. Litonjua AA, Lazarus R, Sparrow D, Demolles D, Weiss ST: **Lung function in type 2 diabetes: the Normative Aging Study.** *Respir Med* 2005, **99**(12):1583–1590.
22. Hickson DA, Liu J, Bidulescu A, Burchfiel CM, Taylor HA, Petrini MF: **Pericardial fat is associated with impaired lung function and a restrictive lung pattern in adults: the Jackson Heart Study.** *Chest* 2011, **140**(6):1567–1573.
23. Regan EA, Hokanson JE, Murphy JR, Make B, Lynch DA, Beaty TH, Curran-Everett D, Silverman EK, Crapo JD: **Genetic epidemiology of COPD (COPDGene) study design.** *COPD* 2010, **7**(1):32–43.
24. Hankinson JL, Odencrantz JR, Fedan KB: **Spirometric reference values from a sample of the general U.S. population.** *Am J Respir Crit Care Med* 1999, **159**(1):179–187.
25. Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR: **MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes.** *Genet Epidemiol* 2010, **34**(8):816–834.
26. Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA, Hurles ME, McVean GA: **A map of human genome variation from population-scale sequencing.** *Nature* 2010, **467**(7319):1061–1073.
27. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC: **PLINK: a tool set for whole-genome association and population-based linkage analyses.** *Am J Hum Genet* 2007, **81**(3):559–575.
28. Willer CJ, Li Y, Abecasis GR: **METAL: fast and efficient meta-analysis of genomewide association scans.** *Bioinformatics* 2010, **26**(17):2190–2191.
29. Castaldi PJ, Dy J, Ross J, Chang Y, Washko GR, Curran-Everett D, Williams A, Lynch DA, Make BJ, Crapo JD, Bowler RP, Regan EA, Hokanson JE, Kinney GL, Han MK, Soler X, Ramsdell JW, Barr RG, Foreman M, van Beek E, Casaburi R, Criner GJ, Lutz SM, Rennard SI, Santorico S, Sciruba FC, DeMeo DL, Hersh CP, Silverman EK, Cho MH: **Cluster analysis in the COPDGene study identifies subtypes of smokers with distinct patterns of airway disease and emphysema.** *Thorax* 2014, **69**(5):415–422.
30. R Core Team: **R: A language and environment for statistical computing.** <http://www.R-project.org/>.
31. Ford ES, Mannino DM, Wheaton AG, Giles WH, Presley-Cantrell L, Croft JB: **Trends in the prevalence of obstructive and restrictive lung function among adults in the United States: findings from the National Health and Nutrition Examination surveys from 1988–1994 to 2007–2010.** *Chest* 2013, **143**(5):1395–1406.
32. Jones RL, Nzekwu MM: **The effects of body mass index on lung volumes.** *Chest* 2006, **130**(3):827–833.
33. **Prevalence of overweight, obesity, and extreme obesity among adults: united states, trends 1960–1962 through 2009–2010.** http://www.cdc.gov/nchs/data/hestat/obesity_adult_09_10/obesity_adult_09_10.htm.
34. Ford ES, Mannino DM, Zhao G, Li C, Croft JB: **Changes in mortality among US adults with COPD in two national cohorts recruited from 1971–1975 and 1988–1994.** *Chest* 2012, **141**(1):101–110.
35. Huseby JS, Petersen D: **Pulmonary function in Klinefelter's syndrome.** *Chest* 1981, **80**(1):31–33.
36. Morales P, Furest I, Marco V, Macian V, Moreno B, Jimenez-Cruz JF: **Pathogenesis of the lung in restrictive defects of Klinefelter's syndrome.** *Chest* 1992, **102**(5):1550–1552.
37. Varkey B, Funahashi A: **Restrictive defect in Klinefelter's syndrome.** *Chest* 1982, **82**(1):132.
38. Daly JJ, Hunter H, Rickards DF: **Klinefelter's syndrome and pulmonary disease.** *Am Rev Respir Dis* 1968, **98**(4):717–719.
39. Rohde R: **Klinefelter's syndrome with pulmonary disease and other disorders.** *Lancet* 1964, **2**(7351):149–150.
40. Bojesen A, Juul S, Birkebaek N, Gravholt CH: **Increased mortality in Klinefelter syndrome.** *J Clin Endocrinol Metab* 2004, **89**(8):3830–3834.
41. Swerdlow AJ, Higgins CD, Schoemaker MJ, Wright AF, Jacobs PA: **Mortality in patients with Klinefelter syndrome in Britain: a cohort study.** *J Clin Endocrinol Metab* 2005, **90**(12):6516–6522.
42. Yamada K, Nomura N, Yamano A, Yamada Y, Wakamatsu N: **Identification and characterization of splicing variants of PLEKHA5 (Plekha5) during brain development.** *Gene* 2011, **492**(1):270–275.
43. Levy D, Ehret GB, Rice K, Verwoert GC, Launer LJ, Dehghan A, Glazer NL, Morrison AC, Johnson AD, Aspelund T, Levy D, Ehret GB, Rice K, Verwoert GC, Launer LJ, Dehghan A, Glazer NL, Morrison AC, Johnson AD, Aspelund T, Aulchenko Y, Lumley T, Kottgen A, Vasan RS, Rivadeneira F, Eiriksdottir G, Guo X, Arking DE, Mitchell GF, Mattace-Raso FU, Smith AV, Taylor K, Scharpf RB, Hwang SJ, Sijbrands EJ, Bis J, Harris TB, Ganesh SK, O'Donnell CJ, Hofman A, Rotter JI, et al: **Genome-wide association study of blood pressure and hypertension.** *Nat Genet* 2009, **41**(6):677–687.
44. Wain LV, Verwoert GC, O'Reilly PF, Shi G, Johnson T, Johnson AD, Bochud M, Rice KM, Henneman P, Smith AV, Ehret GB, Amin N, Larson MG, Mooser V, Hadley D, Dorr M, Bis JC, Aspelund T, Esko T, Janssens AC, Zhao JH, Heath S, Laan M, Fu J, Pistis G, Luan J, Arora P, Lucas G, Pirastu N, Pichler I, et al: **Genome-wide association study identifies six new loci influencing pulse pressure and mean arterial pressure.** *Nat Genet* 2011, **43**(10):1005–1011.
45. Burashnikov E, Pfeiffer R, Barajas-Martinez H, Delpon E, Hu D, Desai M, Borggrefe M, Haissaguerre M, Kanter R, Pollevick GD, Guerschicoff A, Laino R, Marieb M, Nademanee K, Nam GB, Robles R, Schimpf R, Stapleton DD, Viskin S, Winters S, Wolpert C, Simmern S, Veltmann C, Antzelevitch C: **Mutations in the cardiac L-type calcium channel associated with inherited J-wave syndromes and sudden cardiac death.** *Heart Rhythm* 2010, **7**(12):1872–1882.
46. Cordeiro JM, Marieb M, Pfeiffer R, Calloe K, Burashnikov E, Antzelevitch C: **Accelerated inactivation of the L-type calcium current due to a mutation in CACNB2b underlies Brugada syndrome.** *J Mol Cell Cardiol* 2009, **46**(5):695–703.
47. Risaard MS, Jabbari R, Refsgaard L, Holst AG, Haunso S, Sadjdieh A, Winkel BG, Olesen MS, Tfelt-Hansen J: **High prevalence of genetic variants previously associated with Brugada syndrome in new exome data.** *Clin Genet* 2013, **84**(5):489–495.
48. Berger KI, Reibman J, Oppenheimer BW, Vlahos I, Harrison D, Goldring RM: **Lessons from the World Trade Center disaster: airway disease presenting as restrictive dysfunction.** *Chest* 2013, **144**(1):249–257.

doi:10.1186/s12931-014-0089-y
Cite this article as: Wan et al.: Epidemiology, genetics, and subtyping of preserved ratio impaired spirometry (PRISm) in COPDGene. *Respiratory Research* 2014 15:89.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

