Online-only supplement

Deep Neural Network Analyses of Spirometry for Structural Phenotyping of Chronic Obstructive Pulmonary Disease

Sandeep Bodduluri, Arie Nakhmani, Joseph M. Reinhardt, Carla G. Wilson, Merry-Lynn McDonald, Ramaraju Rudraraju, Byron C. Jaeger, Nirav R. Bhakta, Peter J. Castaldi, Frank C. Sciurba, Chengcui Zhang, Purushotham V. Bangalore, and Surya P. Bhatt.

STROBE Statement—Checklist of items that should be included in reports of ***cross-sectional studies***

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2-3
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4-5
Objectives	3	State specific objectives, including any prespecified hypotheses	5
Methods			
Study design	4	Present key elements of study design early in the paper	11
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	11-12
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants	12
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	12
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	*13-14*
Bias	9	Describe any efforts to address potential sources of bias	14
Study size	10	Explain how the study size was arrived at	11, Fig 1
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	11-15
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	15
		(*b*) Describe any methods used to examine subgroups and interactions	15
		(*c*) Explain how missing data were addressed	N/A
		(*d*) If applicable, describe analytical methods taking account of sampling strategy	N/A
		(*e*) Describe any sensitivity analyses	N/A
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	5
		(b) Give reasons for non-participation at each stage	5
		(c) Consider use of a flow diagram	Fig 1
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	5
		(b) Indicate number of participants with missing data for each variable of interest	5
Outcome data	15*	Report numbers of outcome events or summary measures	5

Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	6-7
		(*b*) Report category boundaries when continuous variables were categorized	6-7
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	N/A
Discussion			
Key results	18	Summarise key results with reference to study objectives	7
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	10
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	7-11
Generalisability	21	Discuss the generalisability (external validity) of the study results	10
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	3

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

Additional Results

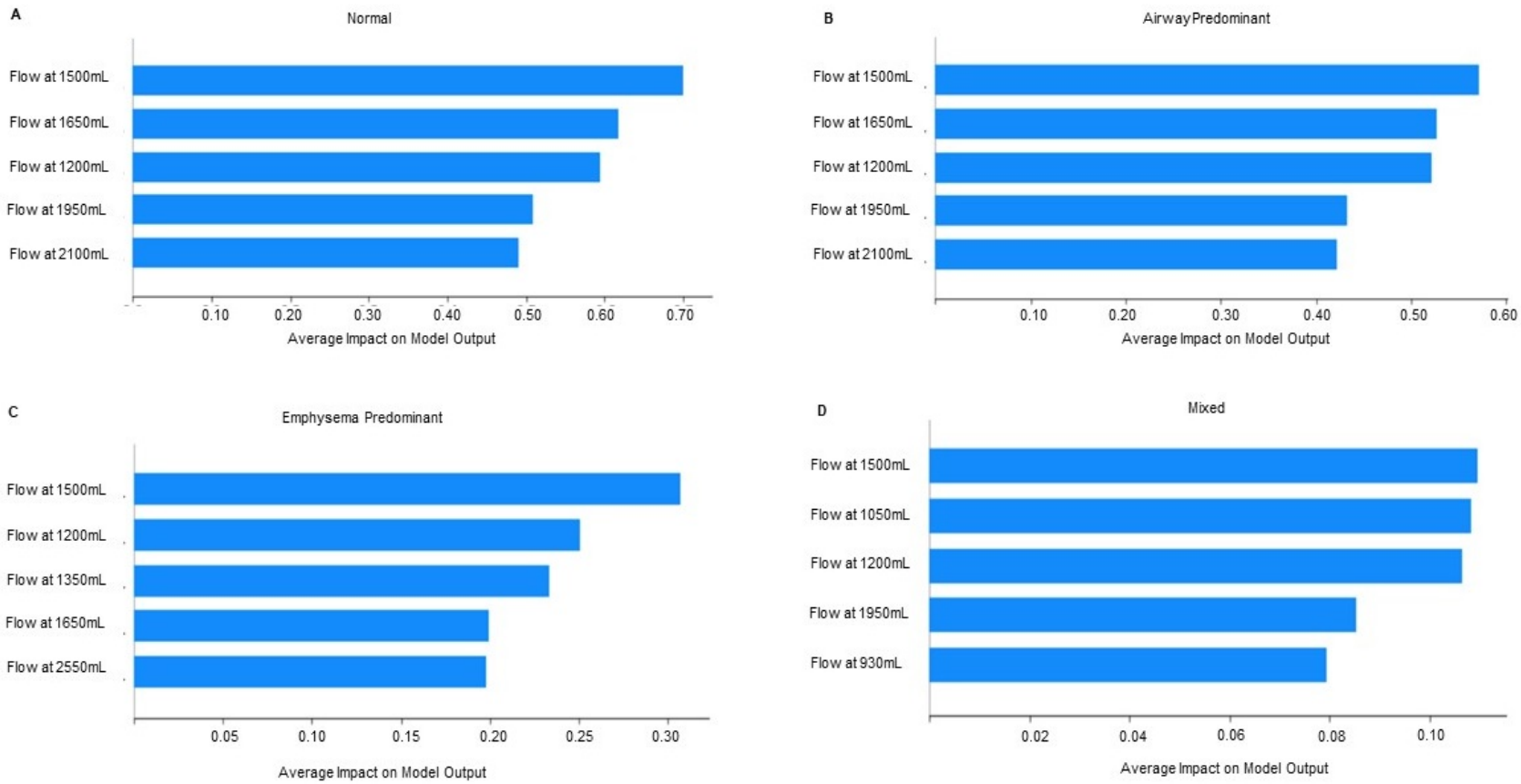

Figure S1: Plots showing the best five features (flow values at specific volumes) on the expiratory flow-volume curve for the prediction of each structural phenotype. Emphysema >5% and Pi10 >median were used to define CT phenotypes. Feature importance was estimated from the results of the fully convolutional network model by SHapley Additive exPlanations (SHAP).

Figure S2 Top Panel

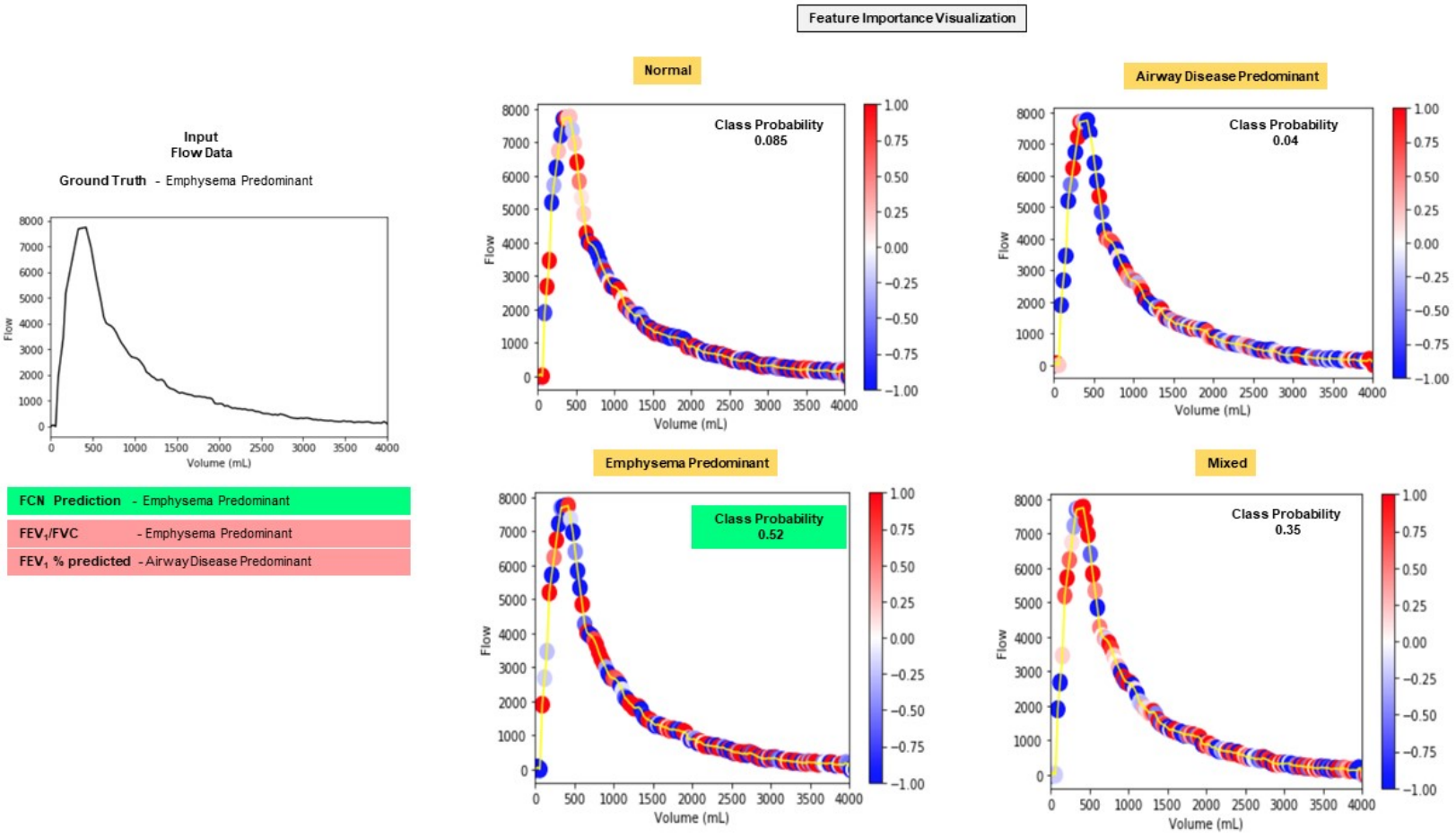

Figure S2 Bottom Panel

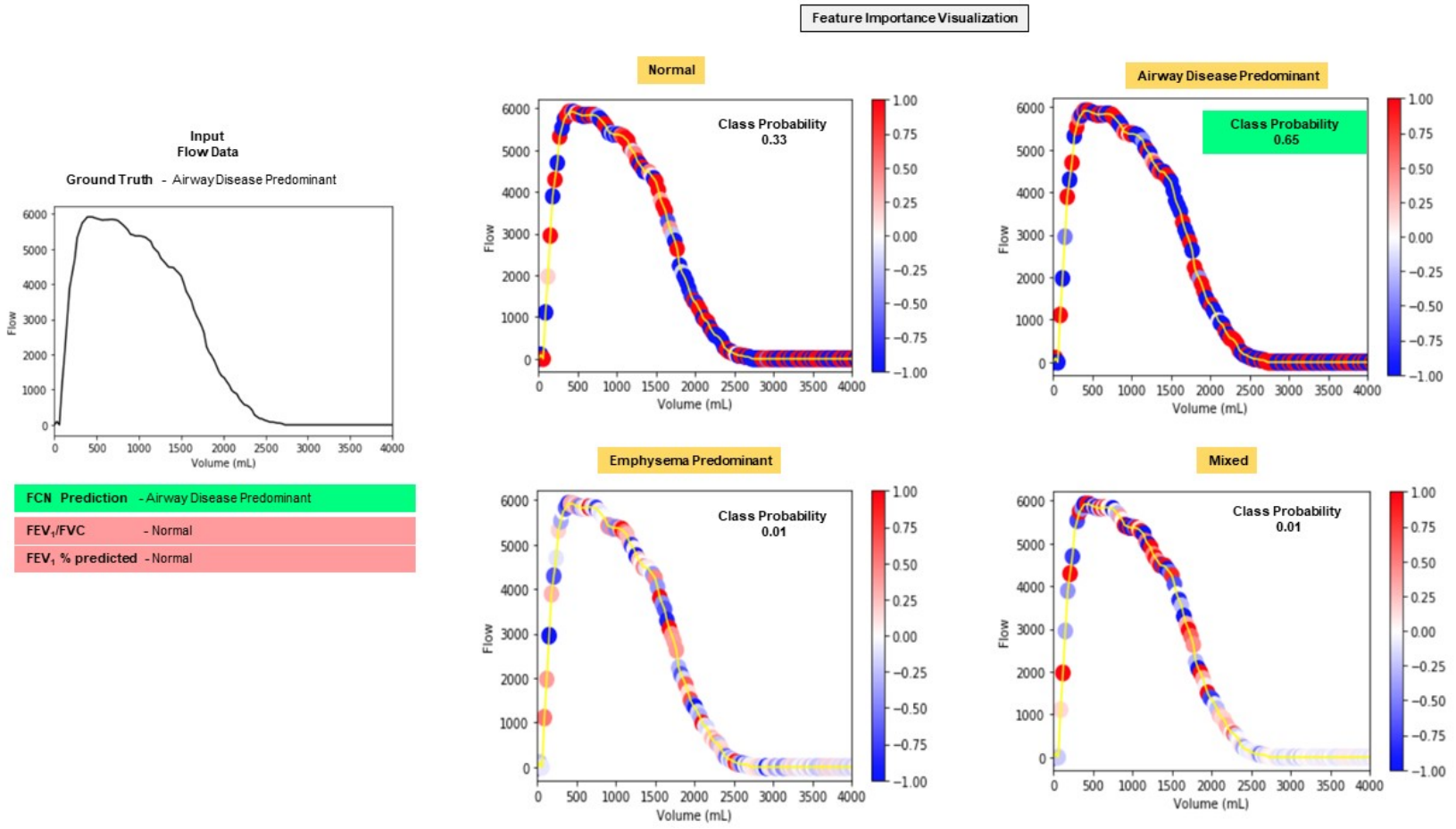

Figure S2: Visualization of SHapley Additive exPlanations (SHAP) values (feature importance) given to each data point on the expiratory flow-volume curve by FCN for the prediction of representative structural phenotypes. Emphysema >5% and Pi10 >median were used to define CT phenotypes.

Top Panel shows Airway-predominant COPD and **Bottom Panel** shows Emphysema-predominant COPD. The feature importance is represented by normalized SHAP values where SHAP value in red represents high impact on the model output and SHAP value in blue represents low impact on the model output. Figure 4a shows a representative flow-volume curve from an individual with airway-predominant disease where the FEV_1/FVC and FEV_1 %predicted measurements misclassified the individual as normal. The FCN, on the other hand, correctly classified the flow-volume curve into the airway-predominant category. Figure 4b shows a flow-volume curve for an individual with emphysema predominant disease where the FEV_1/FVC measurements misclassified the individual as having mixed phenotype and FEV_1 %predicted measurement as airway predominant phenotype. FCN was able to accurately classify this individual.

FEV_1 = Forced expiratory volume in the first second. FVC = Forced vital capacity.

Additional Methods

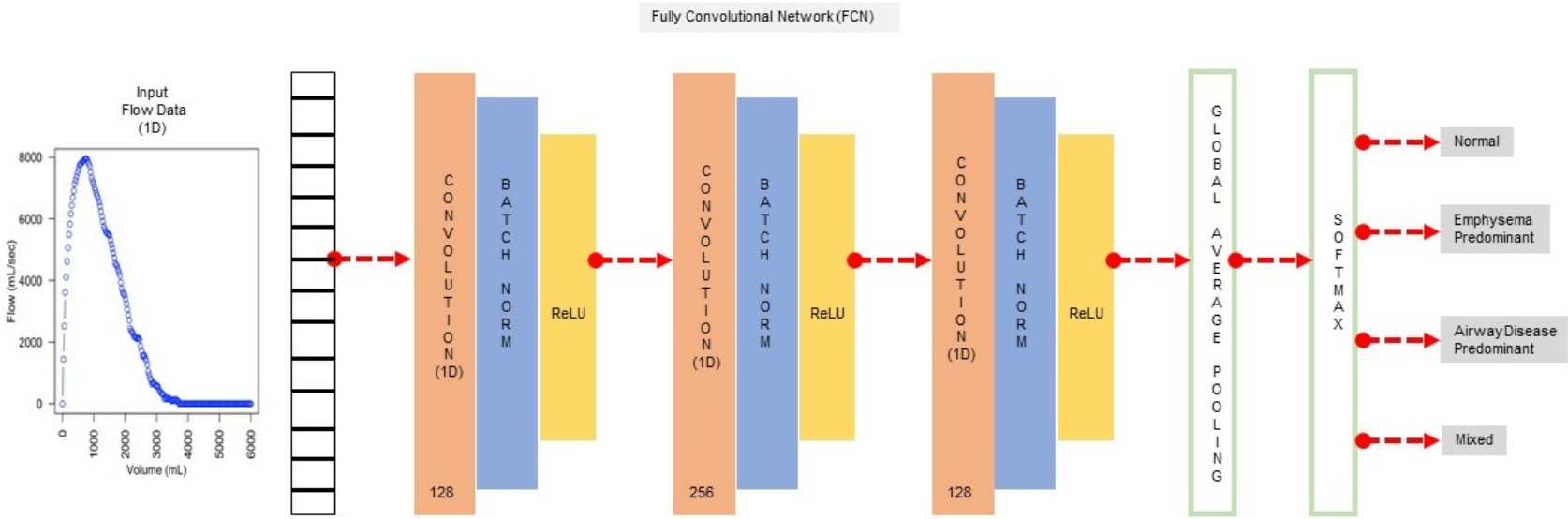

Figure S3: Architecture of Fully Convolutional Network (FCN). 'BATCH NORM' represents batch normalization layer. 'ReLU' represents rectified linear unit activation layers. 'SOFTMAX' represents the final dense 4-class classification layer.

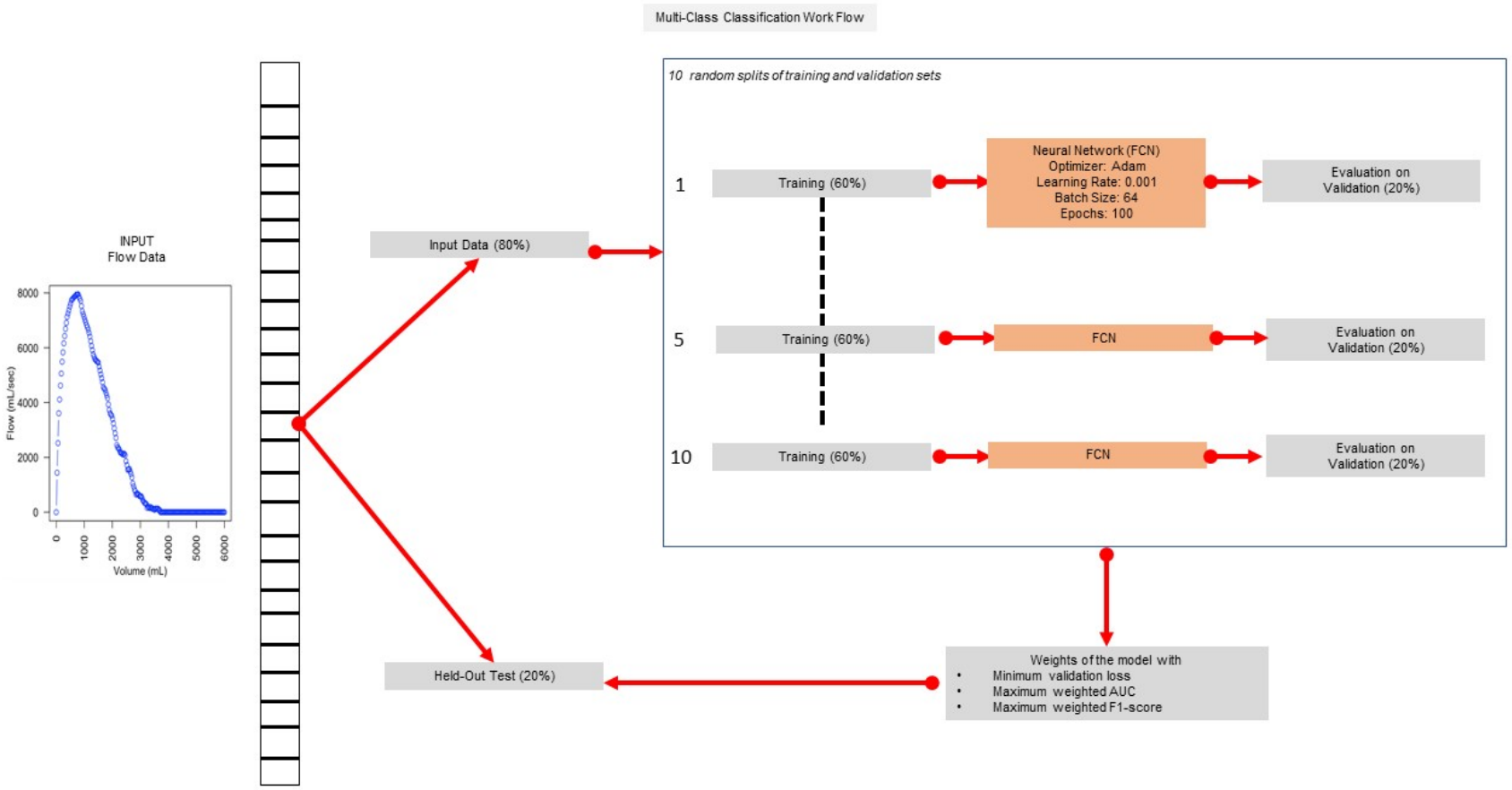

Figure S4: Multi-class classification workflow. The input sequence was divided into 20% held out test dataset and remaining 80% (60% training and 20% validation) was used for training the neural network. The training procedure was repeated for 10 random splits of

training and validation where the weights of the model with minimum validation loss were selected for evaluation in the test dataset. Weighted area under the curve (AUC) and weighted F1-score were used as performance evaluation measures.

Feature Importance

We computed SHapley Additive exPlanation (SHAP) values to estimate the approximate contribution of each variable to the FCN's predictions. SHAP is a unified approach that connects game theory with local explanations, where each player is assigned a weight (SHAP value) representing the player's contribution to the team (i.e., a predictor's contribution to a model's predicted value).[46] SHAP values are obtained by retraining the model over all possible combinations of features and measuring how changes in the feature values impact the model's predictions. We defined variable importance as the absolute SHAP value, i.e., variables that contribute the largest amount to the model's predictions are the most important. Using this definition, we identified the five most important predictors for each phenotype. The SHAP values of the most important predictors were overlaid on the expiratory flow-volume curves to visualize important contributors to specific phenotypes. Please note that SHAP provides information on the most important activation features.

We applied SHAP values to identify five features for each structural phenotype that had the greatest impact on the FCNs predictions (**Figure S3**). The feature with the highest SHAP value for most classes was the flow at 1500 ml volume, except for emphysema predominant disease where flow at 1200 ml had the highest SHAP value. The predominant features for airway disease were primarily flows at higher volumes whereas those for emphysema and mixed disease had an admixture of flows at high and low volumes. We also applied SHAP values to generate patient-specific expiratory flow-volume curves that show the relative contribution of different parts of individual's expiratory flow-volume data to their predicted odds of having emphysema-predominant and airway-predominant disease (Figure S4).

Please note that SHAP provides information on the most important activation features. These activation points are displayed for visual illustration.

IRB Approval:

The COPDGene Study was approved by the Institutional Review Boards of all 21 participating clinical centers. Ann Arbor VA Medical Center 2014-060462 (Ann Arbor VA IRB); Baylor College of Medicine H-22209 (IRB for Baylor College of Medicine); Brigham and Women's Hospital 2007P000554 (Partners Human Research Committee); Columbia Univ. Medical Center AAAC9324 (Columbia University IRB); Duke Univ. Medical Center Pro00004464 (Duke University Health System IRB); Johns Hopkins University NA_00011524 (Johns Hopkins Medicine IRB); L.A. Biomedical Research Inst. 12756-03 (John F. Wolf, M.D. Human Subjects Committee); Michael E. DeBakey VAMC H-22202 (Institutional Review Board for Human Subject Research for Baylor College of Medicine and Affiliated Hospitals); Minneapolis VA Medical Center 4128-A (Minneapolis VA Health Care System Minnesota); Health Partners Twin Cities 07-127 (Health Partners IRB); Morehouse School of Medicine 97826 (Morehouse School of Medicine IRB); National Jewish Health 1883a (National Jewish Health IRB); Reliant Medical Group (Fallon) 1441 (Reliant Medical Group IRB); Temple University 21659 (Temple IRB); Univ. of Alabama, Birmingham F070712014 (University of Alabama at Birmingham IRB for Human Use); Univ. of California, San Diego 140070 (UCSD Human Research Protections Program); University of Iowa 200710717 (University of Iowa IRB); University of Michigan HUM00014973 (University of Michigan Medical School IRB); University of Minnesota 0801M24949 (University of Minnesota IRB Human Subjects Committee); University of Pittsburgh #07120059 (University of Pittsburgh IRB); and UTHSC at San Antonio HSC20070644H (UT Health Science Center San Antonio IRB).