Predictors of Diffusing Capacity in Children with Sickle Cell Disease: A Longitudinal Study

Pritish Mondal[1], Vishal Midya[2], Arshjot Khokhar[1], Shyama Sathianathan[1], and Erick Forno[3]

[1]Penn State College of Medicine
[2]Icahn School of Medicine at Mount Sinai
[3]Children's Hospital Pittsburgh

January 1, 2021

Abstract

Rationale: Gas exchange abnormalities in Sickle Cell Disease (SCD) may represent cardiopulmonary deterioration. Identifying predictors of these abnormalities in children with SCD (C-SCD) may help us understand disease progression and develop informed management decisions. Objectives: To identify pulmonary function tests (PFT) and biomarkers of systemic disease severity that are associated with and predict abnormal carbon monoxide diffusing capacity (DLCO) in C-SCD. Methods: We obtained PFT data from 51 C-SCD (115 observations) and 22 controls, and identified predictors of DLCO for further analyses. We formulated a rank list of DLCO predictors based on machine learning algorithms (XGBoost) or linear mixed-effect models and compared estimated DLCO to the measured values. Finally, we evaluated the association between measured and estimated DLCO and clinical outcomes, including SCD crises, pulmonary hypertension, and nocturnal hypoxemia. Results: DLCO and several PFT indices were diminished in C-SCD compared to controls. Both statistical approaches ranked FVC%, neutrophils(%), and FEV25%-75% as the top three predictors of DLCO. XGBoost had superior performance compared to the linear model. Both measured and estimated DLCO demonstrated significant association with SCD severity indicators. DLCO estimated by XGBoost was associated with SCD crises (beta=-0.084 [95%CI -0.134, -0.033]) and with TRJV (beta=-0.009 [-0.017, -0.001]), but not with nocturnal hypoxia (p=0.121). Conclusions: In this cohort of C-CSD, DLCO was associated with PFT estimates representing restrictive lung disease (FVC%), airflow obstruction (FEV25%-75%), and inflammation (neutrophil%). We were able to use these indices to estimate DLCO, and show association with disease outcomes, underscoring the prediction models' clinical relevance.

Introduction:

Sickle cell disease (SCD) is a hemoglobinopathy that leads to a chronic inflammatory state resulting in vasculitis, pulmonary fibrosis, and pulmonary hypertension[1]. Children with SCD (C-SCD) often suffer from impaired gas exchange, primarily due to hemoglobinopathy and related inflammatory pathology[2]. If untreated, gas exchange abnormalities in SCD may result in chronic hypoxemia, cardiopulmonary morbidity, and poor disease outcomes[3]. Chronic hypoxemia in SCD can contribute to the pathophysiology of vaso-occlusive crises (VOC) and acute chest syndrome (ACS)[4], and it may also lead to pulmonary hypertension, which can impact life expectancy in this vulnerable population[5,6]. Quantifying the underlying pathophysiologic changes is not feasible in routine clinical practice, and thus gas exchange impairment could be used as a prognostic indicator of disease severity in SCD[7].

The single-breath technique for estimating carbon monoxide uptake, also known as DLCO, is a widely used gas exchange measurement technique[8]. Chronic airway inflammation in SCD can lead to worsening diffusion capacity[2]; DLCO impairment also depends on the presence of hypoventilation[9], as well as the

degree of anemia[10]. Despite the importance of DLCO in C-SCD, very few studies have been published on diffusion impairment in C-SCD, and there is no available data on the determinants of DLCO in C-SCD other than anemia. Addressing that knowledge gap could help gain further insight into its origins and prevent morbidities related to impaired gas exchange.

Both DLCO and lung volumes have a faster rate of decline in SCD than healthy subjects. While the relationship is likely complex, it could have prognostic significance; however, it has never been studied before. In the non-SCD population, relationships between DLCO and FVC have been used to stratify mortality risk in pulmonary hypertension[11,12]. Since SCD can lead to pulmonary parenchymal disease and be complicated by pulmonary hypertension, the above-mentioned example underscores the importance of studying the predictors of DLCO and their complex interaction.

Anemia is a primary determinant of DLCO in SCD[13,14]. Subjects with low hemoglobin typically have under-estimated DLCO. Therefore, for precise interpretation, DLCO should be adjusted for hemoglobin in C-SCD. Alveolar ventilation (V_A) is also a strong determinant of DLCO, and previous studies have shown an association between airflow obstruction and diffusion impairment in adults[15]. However, there have been no similar studies in C-SCD. We previously demonstrated the utility of impulse oscillometry (IOS) to measure obstructive airway disease in C-SCD[16], but it is still unknown whether airway resistance or reactance is associated with or predicts gas exchange in C-SCD. Thus, the association between DLCO and measures of airflow obstruction including FEV1, FEV1/FVC, $FVC_{25\text{-}75}\%$, and IOS estimates (R5, X5), is a clinically relevant yet relatively unexplored domain. Unlike obstructive airway disease, restrictive lung disease can be a late manifestation in C-SCD[17], and thus measures like total lung capacity (TLC) and vital capacity (VC) could be significant predictors of declining DLCO –which is more evident with advancing age in C-SCD[18].

In this study, we aim to better understand the predictors of DLCO and their relative importance. Our primary objective was to identify PFT indices and biomarkers that are associated with and predict DLCO in these patients and assess their predictive accuracy. Our secondary objective was to determine if estimated DLCO (eDLCO) is associated with clinical outcomes in C-SCD, which would further emphasize the clinical relevance of DLCO.

Methods:

Study population: We completed a retrospective chart review on 140 C-SCD, ages 6-19 years, followed at the Penn State Pediatric Comprehensive SCD clinic between 2010-2020. PFTs (spirometry, IOS, plethysmography, and DLCO) are typically obtained annually along with pertinent laboratory data. We accessed the charts and extracted demographic characteristics, anthropometric measures, PFT data, pertinent laboratory results, and measures of clinical outcomes.

Control group: We identified 22 race-matched children (African American and Hispanic) without SCD from our patient pool, who performed DLCO between 2018-2020, primarily due to dyspnea of unknown origin. Children with pre-existing cardiovascular, hematological, oncological, or pulmonary conditions that could affect DLCO were excluded. Since data on total hemoglobin were unavailable for most control subjects, we compared DLCO adjusted for alveolar ventilation ($DLCO/V_A$) between cases and controls (the rest of the analyses in C-SCD were performed using hemoglobin-adjusted DLCO, as described above).

Predictors of adjusted DLCO: DLCO was adjusted for hemoglobin concentration and age using sex-specific predictive equations and expressed as a percent of predicted (%pred)[19]. We selected the following potential predictors of DLCO: 1) Pulmonary function test estimates: PFT estimates representing obstructive and restrictive airway disease were considered as potential predictors of DLCO. Spirometry data included forced vital capacity (FVC), forced expiratory volume in 1 second (FEV1), FEV1/FVC, and the forced expiratory volume between 25^{th}-$75^{th}$of FVC (FEV25%-75%). Plethysmography data included total lung capacity (TLC), vital capacity (VC), residual volume (RV), and RV/TLC. Spirometry and plethysmography indices were expressed as %pred (FEV1/FVC and RV/TLC were expressed as a percent). NHANES III equations were used to calculate %predicted values. Measures of total airway resistance (R5) and reactance (X5, Fres, and AX) were obtained from the IOS reports and were expressed as %pred using Berdel/Lechtenbörger

equations (except AX, which does not have standard reference values)[20]. Subjects were instructed not to take bronchodilator therapy for at least 12 hours prior to the PFTs. 2) Laboratory values: the degree of anemia and biomarkers of hemolysis (LDH, total bilirubin, reticulocyte count)– is known to be correlated with SCD related complications. Systemic diseases, including liver and renal function abnormalities, also known to affect DLCO. Neutrophilia and renal failure has been reported as major predictors of death in SCD[5]. Thus, we adjusted the study analyses for SCD biomarkers, including a complete blood count (CBC) with differential, fetal hemoglobin (HbF), and lactate dehydrogenase (LDH) levels, along with liver and renal function test results (e-Table 1).

Indicators of disease severity and clinical outcomes: Number of ACS has been reported to have an association with risk of early death as early as age of 10 years in C-SCD[5,21].Clinical severity indicators considered in this study include lifetime number of hospitalizations with ACS and VOC; sleep-related nocturnal hypoxemia (defined as the percent of total sleep time spent with SpO2 of <90%)[22]. Additionally, tricuspid regurgitation jet velocity (TRJV) >2.5 m/s, measured by echocardiography, was considered as a surrogate marker of pulmonary hypertension[23].

Statistical analyses: We used R (version 3.6.1) and SPSS (version 26.0) for data analysis. DLCO estimates falling outside three times the mean Cook's Distance and two-standard deviation of Studentized t-values were considered to be outliers and were excluded from further analysis. We compared case and control groups with Mann-Whitney U-tests, and used Pearson correlations to estimate the association between potential predictors and DLCO. We added with bootstrap correction to Pearson correlation to adjust for non-normality[24].

Prediction models: Variables with a statistically significant association with DLCO were then examined for relative strength estimation using both a machine learning (ML) based tool, XGBoost, and a linear mixed-effects regression model. XGBoost is a precise and resourceful instrument that can be used for any type of regression analysis or ranking of the predictors, as programmed by a user-built prediction model[25]. We hypothesized that the ML tool would perform better compared to linear regression since it can further adjust for non-linear associations. Both models were adjusted for age, sex, race, hemoglobin genotype as they affect pulmonary function in children with SCD [26,27]. Models were adjusted for hydroxyurea, which increases HbF and improves clinical outcomes in SCD[28]; and asthma medications like LABA and ICS, which can significantly elevate PFT estimates. Finally, models were also controlled for the diagnosis of asthma (yes vs. no) since asthma is one of the major comorbidities in C-SCD[29]. We built the XGBoost model based on the five-fold cross-validation (CV) method. Subjects were randomly divided into five equal groups; four of those five groups were selected at a time as training data and the remaining one as test data, and the process was repeated five times. Based on the results, the predictors of DLCO were selected, and the algorithm was built. We discuss further details in **e-Appendix 1** .

Multicollinearity adjustment: We estimated the degree of multicollinearity between different PFT indices based on simple linear regression analyses by including all indices in the model with hemoglobin-adjusted DLCO as the dependent variable. In this analysis, FEV1(%) had a high variance inflation factor (VIF) of 5.92 and was therefore removed from further analyses to minimize multicollinearity and stabilize the standard error estimates[30]; the rest of the predictor variables were included in the final models for both XGBoost and regression analysis.

Ranking of the predictors: Predictors were ranked based on their relative importance determined by "gain" measure in XGBoost and by p-values in the linear mixed model. To quantify the performance of both models in terms of predictive accuracy, we calculated the mean absolute percentage error (MAPE) and correlation coefficient between measured and eDLCO. MAPE values <10% and between 10%-20% are considered as 'excellent' and 'good' forecasting, respectively[31].

Association between DLCO and clinical outcome measures of SCD: To confirm the prognostic importance of DLCO, we analyzed its association with SCD clinical outcomes using linear regression adjusted for age and sex. For the correlational analyses between lifetime events (numbers) of VOC/ACS and DLCO,

we used the median values of DLCO for the subjects with multiple data points. We also conducted correlation analyses between DLCO and other disease severity indicators, including TRJV and the degree of nocturnal hypoxemia. First, we examined measured DLCO, and then we used our prediction models (XGBoost and mixed-effect model) to calculate eDLCO, and further analyzed the association between eDLCO values and outcome measures using linear regression to cross-examine the accuracy and clinical relevance of the prediction models.

Validation of the prediction model: Leave-one-out performance (LOOP) cross-validation was used for the model validation[32]. Using 'LOOP' function, predicted DLCO was estimated for each study subject while the remaining data (111 in this case) was used to train the XGBoost algorithm. This process was repeated to predict DLCO for all of study participants. The forecast's strength was estimated with MAPE and the Pearson correlation coefficient between observed vs. predicted DLCO.

Results:

Case subjects: During the study period, 51 C-SCD performed a total of 115 DLCO measurements (mean of 2.25 DLCO measurements/subject; range: 1-6). The cohort of C-SCD was comprised of 41 African-American and 10 Hispanic children, with 29 and 22 of them being male and female, respectively. HbSS (41/51) was the most common genotype, followed by HbSC (8/51), Hb S/beta-thalassemia (1/51), and Hb-Lepore (1/51) (**Table 1)** . Mean(SD) DLCO was 87.9(17.2)%; there were no differences in DLCO between hemoglobin genotypes (HbSS vs. HbSC; p=0.74) or the two racial/ethnic groups (African American vs. Hispanic; p=0.82). The mean(SD) age and height of the study participants were 13.0(3.7) years and 150.1(17.6) cm, respectively at the time of PFTs. 91%, 50%, and 17% of C-SCD were on hydroxyurea, ICS, and LABA, respectively, around the time of PFTs. The average number of lifetime ACS episodes and VOC was 3.16 (2.63). Results for SCD biomarkers are summarized in

e-Table 1.

Control group: The mean age of the controls was 10.8 (2.9) years (**Table 1**). C-SCD had lower DLCO/VA, FVC(%pred), FEV1(%pred), TLC (%pred), and VC(%pred) compared to controls (**Table 1**). Race/ethnicity and gender distributions were not statistically different between cases and controls **(Table 1)** .

Evaluation of DLCO predictors: The correlations between hemoglobin-adjusted DLCO with PFT estimates, anthropometrics, and biomarkers are presented in **Table 2** . DLCO was moderately and positively correlated with FEV_1(%pred), FVC(%pred), FEV_1/FVC, $FVC_{25\text{-}75}$(%pred), TLC(%pred); and inversely correlated with R5Hz(%pred) and with peripheral blood neutrophilia (either a percent of WBC or as absolute counts). Aspartate Aminotransferase (AST), total bilirubin, and LDH had a positive correlation with hemoglobin-adjusted DLCO. However, those associations were driven by the strong correlation between the laboratory results and hemoglobin (**e-Table 2**)[30], and thus they were not included in further analyses to prevent over-adjustment bias.

Prediction models:

a) ML Tool XGBoost: FVC(%), neutrophil(%), and $FVC_{25\text{-}75}$(%) were the top three predictors, respectively based on 'gain' function(**Table 3**). MAPE for the model was 1.81%, indicating excellent performance.

b) Linear mixed-effects regression analyses: Hydroxyurea, FVC(%), neutrophil(%), and $FVC_{25\text{-}75}$(%) were statistically significant and the top three predictors for adjusted DLCO (**Table 2**). The rest of the predictors analyzed, including FEV1/FVC, R5(%), and TLC(%), were not statistically significant. The regression model reproduced the exact rank list of six predictors as the XGBoost model (**Table 3**). MAPE between measured and eDLCO for the mixed-model was 9.1%, suggesting that XGBoost had superior prediction performance compared to the regression model (**Figure 1**).

Measured and estimated DLCO vs. outcome measures: Measured DLCO was significantly associated with the number of lifetime VOC/ACS events and TRJV (**Table 4**), but not with nocturnal hypoxemia (p=0.13). After adjusting for age and sex, each 1% decrease in DLCO was associated with 0.075 more lifetime

ACS/VOC events (95%CI:-0.120 to -0.030) and 0.009 m/s higher TRJV (95%CI:-0.017 to -0.001). eDLCO, obtained from our predictive models, was also significantly associated with AOC/VOC events and TRJV (**Table 4**): after adjusting for age and sex, each 1% decrease in eDLCO was associated with 0.084-0.102 more lifetime ACS/VOC events (CI:-0.134 to -0.033 for the XGBoost model, and CI:-0.170 to -0.034 for the regression model) and with 0.009-0.014 m/s higher TRJV (CI:-0.017 to -0.001 for XGBoost, and CI:-0.025 to -0.003 for the regression model) **(Table 4).** Overall, results for modeled eDLCO were very close to those obtained with measured DLCO.

Validation of the prediction model: We tested the strength of the prediction model using LOOP method. Estimated DLCO (mean ± SD) was 87.9 ± 17.18 compared to measured DLCO of 87.79 ± 10.87, with good forecasting (MAPE of 17.3%) and significant correlation (r=0.40, p<0.001*) between two groups (figure 2).

Discussion:

In this study in children with sickle-cell disease, we show that PFT estimates representing obstructive airway disease (FEV25%-75%, FEV1/FVC, R5%), restrictive lung disease (FVC%, TLC%), and biomarkers of inflammation (neutrophil%) were associated with DLCO; and that models built based on those variables can calculate "estimated DLCO (e-DLCO)" with precision. Moreover, we demonstrate that DLCO and e-DLCO are significantly associated with worse clinical outcomes, including more frequent ACS/VOC events and evidence of pulmonary hypertension. These results advance our understanding of factors associated with impaired gas exchange in SCD.

Most pediatric SCD centers in the US do not offer a multi-disciplinary clinic, and PFTs –including DLCO– are not routinely obtained in children with SCD. Clinical status can change rapidly in these children, and PFTs along with other biomarkers need to be obtained at close intervals to estimate the correlation among clinical parameters and build a prediction model. Thus, despite the prognostic significance of impaired gas-exchange, DLCO are not always incorporated into a standard of care in C-SCD, and in-depth clinical research on DLCO is rarely conducted.

Children with SCD in our cohort had significantly lower PFTs than their peers without SCD, consistent with previous studies that have reported impaired lung function in SCD[18,33]. On the other hand, we did not find associations between biomarkers of systemic involvement and DLCO, as has been described in adult SCD literature[13]. This could be partially explained by differences in disease severity or progression in adults with SCD compared to younger populations.

Obstructive airway disease is a relatively early phenomenon in SCD lung involvement, and it can be measured both by spirometry and with IOS. We found that FEV25%-75% and FEV1/FVC were positively correlated with DLCO, while R5(%) showed a negative correlation; obstructive airway disease could thus have an association with impaired gas diffusion in children with SCD. One of the novel aspects of this study was our ability to examine the association between IOS estimates and DLCO. Although an association between IOS estimates and DLCO has never been studied in SCD, a negative correlation between airway resistance (measured by IOS) and DLCO has been reported in adult patients with idiopathic pulmonary fibrosis[34]. With age, airway resistance increases[16] and DLCO(%) decreases in C-SCD[18]; thus, the significant inverse correlation between R5(%) and DLCO(%) may represent a parallel decline in gas diffusion and airway obstruction.

Restrictive airway disease is a relatively late phenomenon in youth with SCD[33]. As the disease progresses, lung volumes and DLCO simultaneously decline due to recurrent inflammation, pulmonary hypertension, and eventually pulmonary fibrosis[13,35,36]. The positive correlation we report between DLCO and lung volume indices such as FVC(%) and TLC(%) may indicate that diminished lung volumes further contribute to impaired gas diffusion. Advanced lung disease, either obstructive or restrictive, can affect alveolar ventilation in adults, leading to alterations in DLCO[35]; our results indicate these alterations start early on in children and even in the absence of severe PFT abnormalities.

Recurrent SCD crises lead to parenchymal disease and impaired gas diffusion[18,37]. Neutrophils generate

extracellular traps and stimulate endothelial activation in SCD[38]. Neutrophil activation and other pro-inflammatory pathways in SCD may lead to thromboembolism in the pulmonary microvasculature, triggering VOC[39]. Thus, neutrophilia may indicate disease severity in C-SCD and it is recognized as a major predictor of mortality in SCD[5]. We found that neutrophilia (either absolute neutrophil counts or percent of total white blood cells) were inversely correlated to DLCO, and neutrophil(%) was among the top three predictors of DLCO. Absolute neutrophil counts have been reported to have inverse correlation with DLCO in the general population[40], but to our knowledge, this is the first report correlating neutrophilia with impaired gas exchange in pediatric SCD.

While diffusing capacity is an important biomarker of SCD lung pathology and is associated with clinical outcomes, diffusion limitation and its probable predictors have not been well studied in C-SCD. Using two different statistical approaches, we evaluated PFT and laboratory predictors of DLCO and identified models that were able to accurately calculate eDLCO. eDLCO closely approximated measured values and was also significantly associated with SCD clinical outcomes. While both mixed-effects regression and XGBoost identified the same predictors, the machine learning model achieved higher precision as evident by lower MAPE (1.81% for XGBoost vs. 9.1% for the linear mixed model). While XGBoost had better precision powered by its ability to adjust for non-linear variable interactions, the reproducibility of the rank list by the linear mixed model adds value, reliability, and a more intuitive interpretation of the models. For instance, both models found that FVC had superior predictive ability compared to FVC_{25-75}; these findings are similar to what has been previously reported in adults without SCD[40]. More importantly, we tested the XGBoost algorithm with LOOP and the precision of DLCO prediction was within the accepted range (between 10-20%), which further validates the prediction model[31]. To the best of our knowledge, no previous study has utilized machine-learning tools to estimate DLCO in C-SCD.

The study has several limitations that should be acknowledged. It was a retrospective, single-center study, and thus we cannot evaluate the effect of center-level practices on our results. Since an external cohort was not available to validate the prediction model, further studies will be needed to validate our findings. We lacked racial and genotypical diversity in the study population, although this is probably fairly representative of the SCD population as a whole. Most of the subjects were in their early teens and had stable lung function, and therefore we cannot extrapolate to younger or older ages; the predictor rank list may have been different if young children or in adults with advanced SCD lung disease. At the same time, our study has several strengths. We had repeated longitudinal data for the cohort, including spirometry, lung volumes, and IOS measurements. We used two different statistical approaches; while one was more accurate than the other in estimating DLCO, both selected the same predictors, which included easy to obtain spirometric and laboratory values. Finally, both measured and estimated DLCO were associated with SCD clinical outcomes.

In conclusion, in a cohort of children with SCD, we report several markers associated with impaired gas exchange, including PFT estimates representing restrictive lung disease (FVC%), obstructive airway disease (FEV25%-75%), and inflammation (blood neutrophil%). DLCO was associated with disease severity indicators of SCD, and we were able to use simple predictors to calculate eDLCO, which was significantly associated with disease outcomes. This underscores the clinical relevance of our prediction models and could help to identify children at risk.

References:

1. Haupt HM, Moore GW, Bauer TW, Hutchins GM. The lung in sickle cell disease. Chest 1982;81(3):332-337.

2. Sylvester KP, Patey RA, Kassim Z, Rafferty GF, Rees D, Thein SL, Greenough A. Lung gas transfer in children with sickle cell anaemia. Respiratory physiology & neurobiology 2007;158(1):70-74.

3. Miller GJ, Serjeant GR. An assessment of lung volumes and gas transfer in sickle-cell anaemia. Thorax 1971;26(3):309-315.

4. Setty BY, Stuart MJ, Dampier C, Brodecki D, Allen JL. Hypoxaemia in sickle cell disease: biomarker

modulation and relevance to pathophysiology. The Lancet 2003;362(9394):1450-1455.

5. Platt OS, Brambilla DJ, Rosse WF, Milner PF, Castro O, Steinberg MH, Klug PP. Mortality in sickle cell disease–life expectancy and risk factors for early death. New England Journal of Medicine 1994;330(23):1639-1644.

6. Mondal P, Stefek B, Sinharoy A, Sankoorikal B-J, Abu-Hasan M, Aluquin V. The association of nocturnal hypoxia and an echocardiographic measure of pulmonary hypertension in children with sickle cell disease. Pediatric Research 2019;85(4):506-510.

7. Chambellan A, Dirou S, Ricolleau B, Graveleau J, Masseau A. The value of diffusing capacity for nitric oxide and carbon monoxide in sickle cell disease. Eur Respiratory Soc; 2015.

8. Ogilvie C, Forster R, Blakemore WS, Morton J. A standardized breath holding technique for the clinical measurement of the diffusing capacity of the lung for carbon monoxide. The Journal of clinical investigation 1957;36(1):1-17.

9. Cotton D, Graham B. Effect of ventilation and diffusion nonuniformity on DLCO (exhaled) in a lung model. Journal of Applied Physiology 1980;48(4):648-656.

10. Dinakara P, Blumenthal W, Johnston R, Kauffman L, Solnick P. The effect of anemia on pulmonary diffusing capacity with derivation of a correction equation. American Review of Respiratory Disease 1970;102(6):965-969.

11. Lacedonia D, Carpagnano GE, Galgano G, Schino P, Correale M, Brunetti ND, Ventura V, Di Biase M, Barbaro MPF. Usefulness of FVC/DLCO ratio to stratify the risk of mortality in patients with pulmonary hypertension. Eur Respiratory Soc; 2016.

12. Hsu VM, Chung L, Hummers LK, Wigley F, Simms R, Bolster M, Silver R, Fischer A, Hinchcliff ME, Varga J. Development of pulmonary hypertension in a high-risk population with systemic sclerosis in the Pulmonary Hypertension Assessment and Recognition of Outcomes in Scleroderma (PHAROS) cohort study. 2014. Elsevier. p 55-62.

13. Klings ES, Wyszynski DF, Nolan VG, Steinberg MH. Abnormal pulmonary function in adults with sickle cell anemia. American journal of respiratory and critical care medicine 2006;173(11):1264-1269.

14. Wall MA, Platt OS, Strieder DJ. Lung function in children with sickle cell anemia. American Review of Respiratory Disease 1979;120(1):210-214.

15. Matheson MC, Raven J, Johns DP, Abramson MJ, Walters EH. Associations between reduced diffusing capacity and airflow obstruction in community-based subjects. Respiratory medicine 2007;101(8):1730-1737.

16. Mondal P, Yirinec A, Midya V, Sankoorikal BJ, Smink G, Khokhar A, Abu-Hasan M, Bascom R. Diagnostic value of spirometry vs impulse oscillometry: A comparative study in children with sickle cell disease. Pediatric pulmonology 2019;54(9):1422-1430.

17. Koumbourlis AC, Zar HJ, Hurlet-Jensen A, Goldberg MR. Prevalence and reversibility of lower airway obstruction in children with sickle cell disease. The Journal of pediatrics 2001;138(2):188-192.

18. Biltagi MA, Bediwy AS, Toema O, Saeed NK. Pulmonary Functions in Children and Adolescents with Sickle Cell Disease. Pediatric Pulmonology 2020.

19. Macintyre N, Crapo R, Viegi G, Johnson D, Van der Grinten C, Brusasco V, Burgos F, Casaburi R, Coates A, Enright P. Standardisation of the single-breath determination of carbon monoxide uptake in the lung. European Respiratory Journal 2005;26(4):720-735.

20. Larsen GL, Morgan W, Heldt GP, Mauger DT, Boehmer SJ, Chinchilli VM, Lemanske Jr RF, Martinez F, Strunk RC, Szefler SJ. Impulse oscillometry versus spirometry in a long-term study of controller therapy for pediatric asthma. Journal of Allergy and Clinical Immunology 2009;123(4):861-867. e1.

21. Thomas AN, Pattison C, Serjeant GR. Causes of death in sickle-cell disease in Jamaica. Br Med J (Clin Res Ed) 1982;285(6342):633-635.

22. Hoth KF, Zimmerman ME, Meschede KA, Arnedt JT, Aloia MS. Obstructive sleep apnea. Sleep and Breathing 2013;17(2):811-817.

23. Ambrusko SJ, Gunawardena S, Sakara A, Windsor B, Lanford L, Michelson P, Krishnamurti L. Elevation of tricuspid regurgitant jet velocity, a marker for pulmonary hypertension in children with sickle cell disease. Pediatric blood & cancer 2006;47(7):907-913.

24. Bishara AJ, Hittner JB. Reducing bias and error in the correlation coefficient due to nonnormality. Educational and psychological measurement 2015;75(5):785-804.

25. Chen T, Guestrin C. Xgboost: A scalable tree boosting system. 2016. p 785-794.

26. Willen SM, Cohen R, Rodeghier M, Kirkham F, Redline SS, Rosen C, Kirkby J, DeBaun MR. Age is a predictor of a small decrease in lung function in children with sickle cell anemia. American journal of hematology 2018;93(3):408-415.

27. Chen L, Gong J, Matta E, Morrone K, Manwani D, Rastogi D, De A. Pulmonary disease burden in Hispanic and non-Hispanic children with sickle cell disease. Pediatric Pulmonology 2020.

28. McLaren A, Klingel M, Behera S, Odame I, Kirby-Allen M, Grasemann H. Effect of hydroxyurea therapy on pulmonary function in children with sickle cell anemia. American Journal of Respiratory and Critical Care Medicine 2017;195(5):689-691.

29. Knight-Madden J, Forrester T, Lewis N, Greenough A. Asthma in children with sickle cell disease and its association with acute chest syndrome. Thorax 2005;60(3):206-210.

30. Sheather S. A modern approach to regression with R. Springer Science & Business Media; 2009.

31. Moreno JJM, Pol AP, Abad AS, Blasco BC. Using the R-MAPE index as a resistant measure of forecast accuracy. Psicothema 2013;25(4):500-506.

32. Wong T-T. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. Pattern Recognition 2015;48(9):2839-2846.

33. Sylvester KP, Patey R, Milligan P, Dick M, Rafferty G, Rees D, Thein S, Greenough A. Pulmonary function abnormalities in children with sickle cell disease. Thorax 2004;59(1):67-70.

34. Semenova E, Kameneva M, Tishkov A, Trofimov V, Novikova L. Relationship the impulse oscillometry parameters and the lung damage in idiopathic pulmonary fibrosis patients. Eur Respiratory Soc; 2013.

35. Mehrad B, Burdick MD, Wandersee NJ, Shahir KS, Zhang L, Simpson PM, Strieter RM, Field JJ. Circulating fibrocytes as biomarkers of impaired lung function in adults with sickle cell disease. Blood advances 2017;1(24):2217-2224.

36. Anthi A, Machado RF, Jison ML, Taveira-DaSilva AM, Rubin LJ, Hunter L, Hunter CJ, Coles W, Nichols J, Avila NA. Hemodynamic and functional assessment of patients with sickle cell disease and pulmonary hypertension. American journal of respiratory and critical care medicine 2007;175(12):1272-1279.

37. Sylvester KP, Patey RA, Milligan P, Rafferty GF, Broughton S, Rees D, Thein SL, Greenough A. Impact of acute chest syndrome on lung function of children with sickle cell disease. The Journal of pediatrics 2006;149(1):17-22.

38. Chen G, Zhang D, Fuchs TA, Manwani D, Wagner DD, Frenette PS. Heme-induced neutrophil extracellular traps contribute to the pathogenesis of sickle cell disease. Blood 2014;123(24):3818-3827.

39. Zhang D, Xu C, Manwani D, Frenette PS. Neutrophils, platelets, and inflammatory pathways at the nexus of sickle cell disease pathophysiology. Blood 2016;127(7):801-809.

40. Neas LM, Schwartz J. The determinants of pulmonary diffusing capacity in a national sample of US adults. American journal of respiratory and critical care medicine 1996;153(2):656-664.

Hosted file

`Table 1.pdf` available at https://authorea.com/users/386591/articles/502032-predictors-of-diffusing-capacity-in-children-with-sickle-cell-disease-a-longitudinal-study

Hosted file

`Table 2.pdf` available at https://authorea.com/users/386591/articles/502032-predictors-of-diffusing-capacity-in-children-with-sickle-cell-disease-a-longitudinal-study

Hosted file

`Table 3.pdf` available at https://authorea.com/users/386591/articles/502032-predictors-of-diffusing-capacity-in-children-with-sickle-cell-disease-a-longitudinal-study

Hosted file

`Table 4.pdf` available at https://authorea.com/users/386591/articles/502032-predictors-of-diffusing-capacity-in-children-with-sickle-cell-disease-a-longitudinal-study