

# Analysis of Pulmonary Function Test Results By Using Gaussian Mixture Regression Model

Serdar Abut<sup>1</sup>, Fatih Doğanay<sup>2</sup>, Abdullah Yeşilova<sup>3</sup>, Serap Buğa<sup>4</sup>

<sup>1</sup>Department of Computer Engineering, Siirt University, Siirt, Turkey

<sup>2</sup>Department of Emergency Medicine, Edremit State Hospital, Balıkesir, Turkey

<sup>3</sup>Department of Animal Science, Yüzüncü Yıl University, Van, Turkey

<sup>4</sup>Primary Care Health Center, Balıkesir Provincial Health Directorate, Balıkesir, Turkey

Received: 2021-02-16.

Accepted: 2021-04-04



This work is licensed under a  
Creative Commons Attribution 4.0  
International License

J Clin Med Kaz 2021; 18(2):23-29

Corresponding author:

Serdar Abut.

E-mail: [serdarabut@siirt.edu.tr](mailto:serdarabut@siirt.edu.tr);

ORCID: 0000-0002-6617-6688

## Abstract

**Background:** FEV1/FVC value is used in the diagnosis of obstructive and restrictive diseases of the lung. It is a parameter reported in the literature that it varies according to lung disease as well as weight, age and gender characteristics.

**Objective:** The aim of this study is to investigate the relationship between age, weight, gender and height characteristics and FEV1/FVC value using a heterogeneous population using Gaussian mixture regression method.

**Material and methods:** GMR was used to separate the data into components and to make a parameter estimation for each component. The analysis performed on this model revealed that the patients were divided into 5 optimal groups and that these groups showed a regular transition from obstructive pattern to restrictive pattern.

**Results:** The mean values of the components for FEV1/FVC were found as 50.071 (3.238), 67.034 (1.725), 82.156 (1.329), 93.592 (1.041), 98.466 (0.303), respectively. The effect of the weight on the components in terms of parameter estimation and standard errors of the components was determined as 0.445 (0.129)\*\*; 0.226 (0.053)\*\*; 0.173 (0.053)\*\*; -0.036 (0.026); -0.040 (0.018)\*, respectively.

**Conclusion:** Direct proportional relationship between the patient's weight and the severity of the obstructive pattern, and between the severity of the disease and the age of the patient in both the obstructive and restrictive pattern are explicitly proved. Furthermore, it has been revealed that data sets containing heterogeneity can be analysed by dividing them into sub-components using the GMR model.

**Key words:** Gaussian mixture regression, pulmonary function test, FEV1/FVC, obstructive pattern, restrictive pattern

## Introduction

Pulmonary function test (PFT) is diagnostic for obstructive pulmonary diseases (such as asthma, Chronic Obstructive Pulmonary Disease (COPD), restrictive lung diseases (parenchymal diseases, respiratory muscles, chest wall diseases, pulmonary edema, congestive heart failure) and extra-thoracic airway stenosis (trachea obstruction, vocal cord paralysis) [1]. PFT is a test used to diagnose respiratory system diseases and to monitor the progression of patients with known respiratory diseases [2]. Forced expiratory volume in 1 second (FEV1) and the forced vital capacity (FVC) ratio (FEV1/FVC), is a parameter measured in the PFT, used for the diagnosis of

respiratory system diseases. FEV1/FVC, especially in the separation of obstructive and restrictive diseases of the lung is the most important parameter used PFT [3]. FEV1 (Forced expiratory volume in 1 second) and FVC (Forced vital capacity) are two of the values measured during PFT. FVC is the maximum air volume that can be blown after full inspiration. It can also be considered as the maximum breath volume of the lung (forced inspiratory vital capacity). FEV1 is the volume of air that can be blown in the first 1 second after a full inspiration [4].

When evaluating PFT results, the FEV1/FVC ratio is taken into account first. First, the obstructive-restrictive distinction is performed [5]. The FEV1/FVC ratio, also

called the Tiffeneau-Pinelli index, is a ratio used to differentiate obstructive from restrictive lung diseases. The ratio of FEV1/FVC (FEV1%) in healthy adults should be about 70-80% [4]. It is recommended by the GOLD criteria that patients with FEV1/FVC below 70% should be evaluated as an obstructive disease [6].

In obstructive diseases, airflow from the lungs becomes difficult due to airway obstruction, the expiratory time is extended and the maximum volume of air (FEV1) blown in the first 1 second decreases [7]. FEV1 decreases in obstructive diseases (asthma, COPD, chronic bronchitis, emphysema) due to increased airway resistance in expiration. The main problem in restrictive diseases is fibrosis in the alveolar lung tissue. The compliance of the lung decreases due to this fibrosis. In restrictive diseases, both FEV1 and FVC are reduced. Since the decrease in FVC is greater than the decrease in FEV1, the ratio of FEV1 / FVC increases [8].

In the guidelines of the American Thoracic Society (ATS) published in 1987, it reported that the FEV1/FVC value should be below 75% for the diagnosis of obstructive pulmonary disease [7]. Later, in the guideline published by the British Thoracic Society (BTS) in 1997, it was proposed to use 70% as the upper limit of the FEV1/FVC value for the diagnosis of obstructive pulmonary disease [9]. In a study investigating ways to develop reference ranges on the relationship between spirometric lung function and height and age, the results were modeled in terms of gender, age and height [10]. In a study where two separate expected value formulas for men and women were tried to be developed in order to create equality for two different genders on spirometric test results, age and height variables were also used [11]. In relation to these variables, Swanney et al. stated that it was not appropriate to determine the FEV1/FVC value as a standard cut-off value [12].

Studies on determining the ranges of FEV1/FVC values that should be accepted as normal in adults are available in the literature [12-15].

The aim of this study is to investigate the relationship between age, weight, gender, and height variables with FEV1/FVC values in a heterogeneous population using gaussian mixture regression method and summarize all data with a multivariate statistical method.

## Materials and methods

### PFT equipment

SFT measurements were made with Spirobank II (MIR-Medical International Research USA, Inc., 5462 S. Westridge Drive New Berlin, WI 53151-USA) device. During the time that the data used in the study were obtained, all PFTs were made by the same technician. The calibration of the device was done once a day, following automatic procedures.

### Study Sample

Those who were admitted to a secondary hospital between May 1, 2018 and May 31, 2018 and had a pulmonary function test were included in the study. A dataset was created by scanning PFT records retrospectively. Age, gender, weight and height information on the records, including PFT results, were recorded. Those missing at least one of these information were excluded from the study. After the information of the people with incomplete data was removed from the dataset, the analysis started with the data of the remaining 171 people. FEV1/FVC variable is determined as dependent variable and height, age, weight and gender are determined as independent variables.

## Statistical analysis

While performing statistical analysis, the excess number of variables to be examined makes it difficult to solve by using classical statistical techniques such as clustering and factor analysis. In medical studies, the observed events depend on many factors. At the same time, these factors are interrelated. In order to obtain valid and reliable results, it is recommended that all variables related to the research be participated in the analysis. At this point, the mixture model is a powerful multivariate statistical tool that can be used effectively.

## Finite mixture models

In multivariate statistical analysis techniques such as factor and clustering analysis, multiple variables should be clustered in terms of their effect on dependent variables. As a result of this grouping, it is ideal that both heterogeneity between groups and homogeneity within groups are high [16]. In recent years, mixture model approach has been used extensively in multivariate statistics. Mixture models have two important advantages over clustering and factor analysis. The first advantage is the calculation of which component is likely to include each observation. The second important advantage is that it is possible to calculate parameter estimates for each component with the mixture model. This is because mixture modelling also performs multiple regression analysis specifically for each component after dividing the data into components. Gaussian mixture regression analysis is used if the dependent variable fits normal distribution [17].

Another purpose of using mixture models is minimize the probability of misclassification and determine the populations from which they were sampled. In mixture models, parameters estimate by expectation and maximization (EM) algorithm and Maximum Likelihood (ML) method. Entropy is used for calculating true classification probability in mixture modelling [16]. Akaike and Bayesian information Criteria metrics are used for optimum component number [18].

## Theoretical background

Component densities  $f_i(y_j)$  stated to apply to some parametric distributions. The component densities  $f_i(y_j)$  are described as  $f_i(y_j; \theta_i)$ , where  $\theta_i$  is the vector of unknown parameters. This vector  $\theta_i$  is hypothesized form for the  $i$ th component density in the mixture. Mixture density can be formulated as

$$f(y_j; \psi) = \sum_{i=1}^g \pi_i f_i(y_j; \theta_i) \quad (1)$$

where vector  $\psi$ , involving all unknown parameters in the mixture model can be define as

$$\psi = (\pi_1, \pi_2, \dots, \pi_{g-1}, \xi^T)^T \quad (2)$$

In Eq.2 vector  $\xi$ , contain all the parameters in  $\theta_1, \theta_2, \dots, \theta_g$

Parametric form of mixture model written as

$$f(y_j; \Psi) = \sum_{i=1}^g \pi_i f_i(y_j; \theta_i) \quad (4)$$

where observed random sample  $y = (y_1^T, \dots, y_n^T)^T$ . The loglikelihood for  $\Psi$  that can be form the observed data is given by

$$\log L(\Psi) = \sum_{j=1}^n \log f(y_j; \Psi) = \sum_{j=1}^n \log \left\{ \sum_{i=1}^g \pi_i f_i(y_j, \theta_i) \right\} \quad (5)$$

Likelihood equation written as;

$$\partial \log L(\Psi) / \partial \Psi = 0 \quad (6)$$

EM algorithm trying to solve Eq.6 using iterative methods and Eq.5.

## Model selection

In the mixture model approach, AIC and BIC model fit statistics are widely used to determine how many sub-grouped models are homogeneous within each other [19-21]. The model with the smallest AIC and BIC values is considered to be the best representation of the data set [22]. However, in determining the model that best fits the data set, entropy, which has the possibility of correct classification, is also widely used. The entropy criterion can be considered as the best model in which component model is higher [23].

AIC and BIC were used to determine the optimum number of components. AIC, selects model that can minimize Eq.7.

$$AIC = -2 \log L(\hat{\Psi}) + 2d \quad (7)$$

After Akaike, new approaches to the use of AIC to select the number of components in a mixture [23-25].

$$\log p(y) = \log L(\hat{\Psi}) + \log p(\hat{\Psi}) - \frac{1}{2} \log |I(\hat{\Psi}; y)| + \frac{1}{2} d \log(2\pi) \quad (8)$$

Schwarz [26] suggested that Equation 8 omits the term  $O(1)$ ;

$$|I(\hat{\Psi}; y)| = O(n^d) \quad (9)$$

And Bayesian information criterion (BIC) for the number of components  $g$  in the mixture model, to be obtained as the smallest in the model selection, with a negative log penalty multiplied by two and written as;

$$-2 \log L(\Psi) + d \log n \quad (10)$$

## Results

Surface plots of independent variables height – weight, age – height and age – weight versus FEV1/FVC, are shown in Figure 1, Figure 2 and Figure 3 respectively. These plots created by biharmonic interpolation method. These figures show that the response variable FEV1/FVC have heterogeneous structure versus the independent variables.

These heterogeneous structures were analyzed by homogenous fractionation with using gaussian mixture regression model. Akaike and Bayesian Information Criteria which used to determine how many components and entropy correct classification rates are given in Table 1 and Figure 4.

As a result of the 5 component model obtained using the normal mixture regression model, the number and mixing probability of the individuals entering each component were shown as Table 1. The highest number of individuals was in component 5 and the lowest number of individuals was in component 1 (Figure 5, Figure 7).

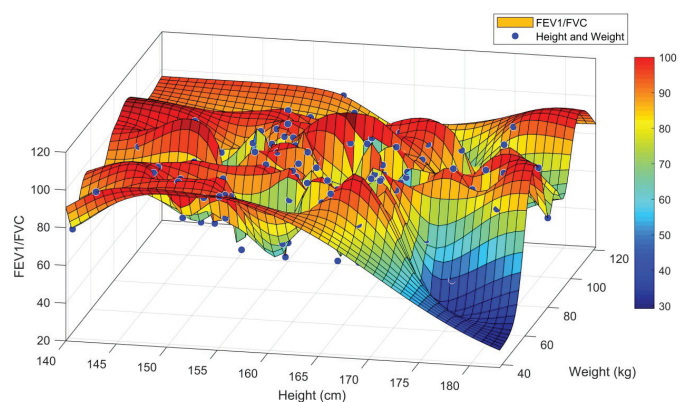


Figure 1. Surface plots of FEV<sub>1</sub>/FVC versus height and weight variables

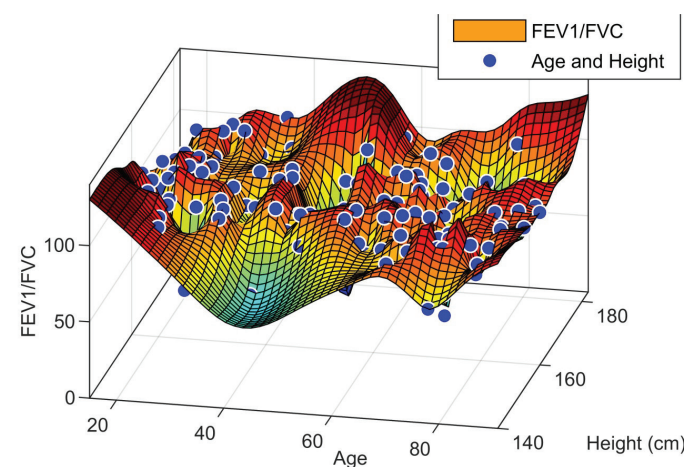


Figure 2. Surface plots of FEV<sub>1</sub>/FVC versus age and height variables

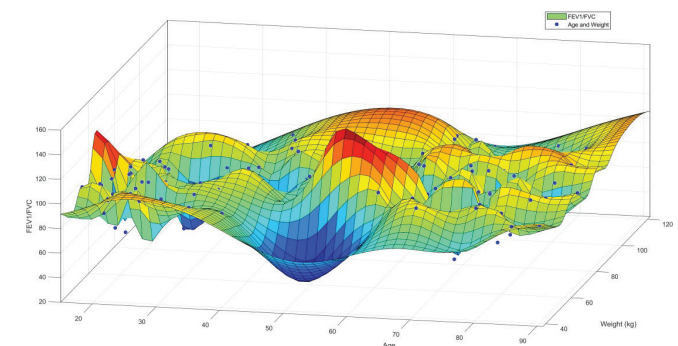


Figure 3. Surface plots of FEV<sub>1</sub>/FVC versus age and weight variables

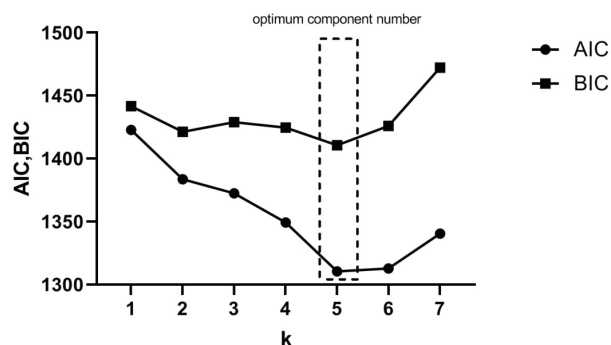


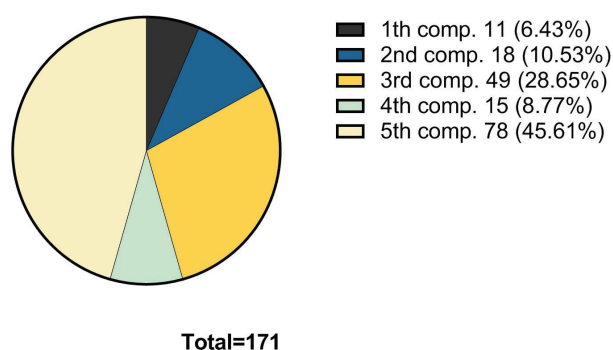
Figure 4. AIC and BIC values versus component number (k)



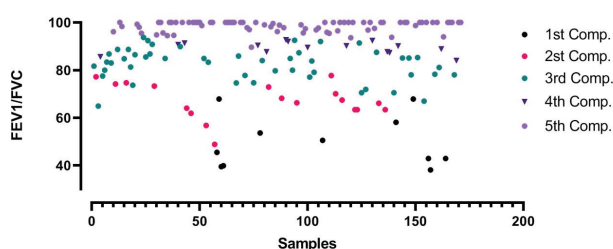
Table 1

Model selection criteria for components

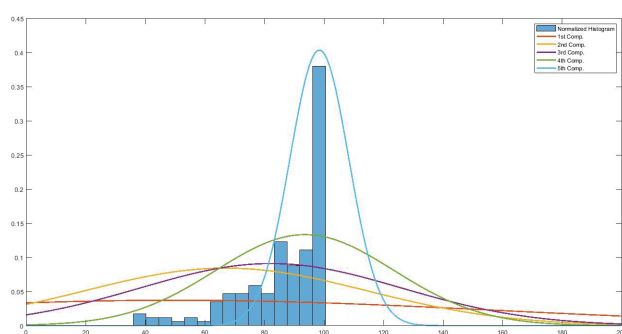
Components	AIC	Model Information Criteria	
		BIC	Entropy (%)
Model with one component	1422.759	1441.609	87.0
Model with two component	1383.531	1421.231	84.5
Model with three component	1372.400	1428.950	89.3
Model with four component	1349.211	1424.611	90.8
Model with five component	1310.366	1410.616	92.5
Model with six component	1312.757	1425.857	91.9
Model with seven component	1340.366	1472.316	86.2



**Figure 5.** Obtained class sizes and mixing probabilities for components with closest class membership probabilities



**Figure 6.** Distribution of dependent variable according to five components



**Figure 7.** Prior conditional probability density functions of gaussian components

In the model with 5 components, which are the best model in which the data set is distributed (Figure 6), the parameter estimates are given in Table 2.

In the first component, the effect of weight independent variable on FEV1/FVC dependent variable was statistically significant ( $p < 0.01$ ); height, age and gender variables were statistically insignificant ( $p > 0.05$ ). In the second component, the effects of weight and age independent variables on FEV1/FVC dependent variable were statistically significant ( $p < 0.01$ ),

the effect of height and gender independent variables was found to be statistically insignificant ( $p > 0.05$ ). In the third component, the coefficient of height independent variable on FEV1/FVC dependent variable was found to be statistically insignificant ( $p > 0.05$ ); however, the effects of age, gender ( $p < 0.05$ ) and weight ( $p < 0.01$ ) independent variables were statistically significant. In the fourth component, gender independent variable on FEV1/FVC dependent variable were statistically significant ( $p < 0.05$ ) and the effects of weight, height and age independent variables on FEV1/FVC dependent variable were found to be statistically insignificant ( $p > 0.05$ ). Finally, in the fifth component, the effect of weight independent variable on FEV1/FVC dependent variable was statistically significant ( $p < 0.01$ ); height, gender and age independent variables were statistically insignificant ( $p > 0.05$ ) (Table 2).

Mean values of FEV1/FVC dependent variable and estimated means and standard errors of gender, height, weight and age variables are shown in Table 3.

The normal distribution density functions of the five different components formed according to the obtained regression parameters are given in Figure 7. Considering the class sizes obtained for the components with the closest class membership (probability) grades, it was found that the mixing ratio of the fifth component had the highest mixing ratio as 45.61%.

## Discussion

FEV1/FVC value is expected to decrease below 70% in obstructive pulmonary diseases, while it is expected to be normal or increase above 75% in restrictive lung diseases [12, 27-29]. FEV1/FVC value for the diagnosis of COPD is expected to be below 70% [28]. In obstructive disorders, there is a disproportionate decrease in the maximum volume of air exhaled from the lung. This represents airway contraction during exhalation and is defined by the FEV1/FVC ratio being less than 70% of the source values [29].

In this study, it was found that AIC and BIC values decreased to 5 component model and then increased. Thus, AIC and BIC with the smallest statistics and entropy value with high accurate classification ratio were obtained in 5 component model (Figure 4). The entropy criterion was obtained as 92.5% in the 5 component model selected as the most fitted model. It is observed that variable FEV1/FVC grows with component index. When components are examined, a regular and significant transition from low to high values for FEV1/FVC variable is observed.

The first component stands out as the component in which the obstructive pattern is most prominent. Considering that the average of the weight variable in component 1 is higher than the average in other components, it can be said that the FEV1/FVC is inversely proportional to weight in patients with low FEV1/FVC. The relationship between obesity and COPD has not been clearly determined yet. In 2014 García-Rio et al. reported that obesity was more prevalent in patients diagnosed with COPD compared to the general population [28]. There are studies reporting improvement in PFT results after weight loss in patients undergoing bariatric surgery [29]. In a study comparing the FEV1/FVC value of obese children compared to normal-weight children, it was reported that the FEV1/FVC value decreased in direct proportion to the increase in BMI (Body Mass Index) [30]. A study published in 2015 reported that when a fixed FEV1/FVC ratio of 0.70 used, it resulted in under-diagnosis of COPD among overweight patients [31]. A review by Dixon et al. has been reported that obesity may cause some

Table 2

Parameter estimation and standard errors of components

Component	Intercept $\beta_0$	Mean of FEV <sub>1</sub> /FVC	Independent variables Estimated Coefficient (Standard Error)			
			Gender (1:Male, 2:Female)	Height (cm)	Weight (kg)	Age (years)
1	-101.05	50.071	8.153 (7.746)	0.626 (0.552)	0.445 (0.129)**	-0.006 (0.179)
2	101.93	67.034	3.984 (3.300)	-0.295 (0.166)	0.226 (0.053)**	-0.226 (0.035)**
3	42.73	82.156	9.144 (4.571)*	0.137 (0.230)	0.173 (0.053)**	-0.188 (0.074)*
4	95.60	93.592	7.740 (2.207)*	-0.074 (0.107)	-0.036 (0.026)	0.019 (0.025)
5	95.06	98.466	-0.467 (0.558)	-0.037 (0.028)	-0.040 (0.018)*	0.019 (0.010)

\* coefficient is statistically significant at the alpha=0.05 levels

\*\* coefficient is statistically significant at the alpha=0.01 levels

Table 3

Means of variables according to components

		Independent variables Estimated Mean (Standart Error)				
Component	Mean of FEV1/FVC	Gender (Count, %)		Height (cm)	Wight (kg)	Age (years)
		Male	Female			
1	50.071(3.238)	4 (36.4%)	7 (63.6%)	164.18 (2.717)	78.862 (4.540)	47.643 (6.859)
2	67.034(1.725)	7 (38.9%)	11 (61.1%)	164.80 (1.806)	74.775 (3.713)	42.548 (5.036)
3	82.156(1.329)	17 (34.7%)	32 (65.3%)	163.816 (1.322)	70.099 (2.167)	54.843 (3.009)
4	93.592(1.041)	15 (100%)	0 (0%)	162.667 (1.610)	75.295 (3.181)	55.521 (4.084)
5	98.466(0.303)	30 (38.5%)	48 (61.5%)	162.798 (1.229)	71.341 (1.970)	53.221 (3.014)

changes in the mechanical functions of the lungs and chest wall. In addition, it was concluded that obesity affects lung function through an inflammatory pathway by increasing inflammatory cytokines [32]. It should be noted that the relationship between obesity and FEV<sub>1</sub>/FVC is not only related to weighted increase, but also the disease-causing effects.

When the components are ranked from small to large in terms of FEV<sub>1</sub>/FVC scores, the second component with a value of 67.034 FEV<sub>1</sub>/FVC is in the second place. In this component, the effects of weight and age variables on FEV<sub>1</sub>/FVC were found statistically significant ( $p < 0.01$ ). It was determined that FEV<sub>1</sub>/FVC value decreased -0.226 times when age increased by one unit and FEV<sub>1</sub>/FVC value increased by 0.226 times when weight increased. The second component population is a component composed of individuals with an FEV<sub>1</sub>/FVC value close to 70%. The minimum and maximum FEV<sub>1</sub>/FVC values of this component are values that converge to 70%.

When the effect of age independent variable on FEV<sub>1</sub>/FVC is analyzed, it is determined that the coefficient of the age variable is negative in the components where the average value of FEV<sub>1</sub>/FVC is low. However, it was found that the coefficient for the age variable turned positive in components where FEV<sub>1</sub>/FVC was close to the 70% threshold value (Table 2). As age increases in patients at this component, FEV<sub>1</sub>/FVC decreases more clearly than other components. It is concluded that FEV<sub>1</sub>/FVC value significantly increases with increasing age in this component. In a study conducted by collecting data between 1998 and 2009 in the United States, it was concluded that the prevalence of COPD increases with age for both men and women [33].

Changes such as a significant decrease in lung elasticity, chest wall stiffness and loss of respiratory muscle strength are the physiological effects that occur in the respiratory system

with aging [34-36]. These physiological changes were seen as the cause of low FEV<sub>1</sub>/FVC in the elderly population [12].

The effect of the height independent variable on the FEV<sub>1</sub>/FVC was found insignificant in all components.

Considering the gender independent variable, the effect on the FEV<sub>1</sub>/FVC dependent variable was found statistically significant only in the third and fourth components. It was concluded that in cases where FEV<sub>1</sub>/FVC values were slightly higher or slightly lower, the male sex was predisposed to form groups with moderately low FEV<sub>1</sub>/FVC, while the female sex was predisposed to form groups with moderately high FEV<sub>1</sub>/FVC. The third component stands out as the component in which the gender variable is statistically most significant. The relationship between COPD and gender has been investigated since 1987 [37]. McHugh et al. also reported that FEV<sub>1</sub>/FVC values differ significantly between men and women [38].

## Limitations

Retrospective studies have some limitations due to their nature as the data are scanned retrospectively. Our study was conducted with data obtained from 171 patients, and the sample size is relatively small. Prospective studies with larger samples are needed to verify the results of this study. Since we did not investigate the effect of smoking on FEV<sub>1</sub>/FVC, we did not use smoking history in our study. We think that it would be useful to compare the results of this study with the results of studies with large populations that are homogenized and free from other effects.

## Conclusion

In this study, the heterogeneity in the dataset is tried to be understood with the mixture regression model. Using the

mixture regression model gives useful results in the analysis of dependent variables that interact with multiple independent variables, both in determining the existence and power of the relationships. We can say that the relationship between FEV1/FVC value and weight is greater than with other variables, in terms of the characteristics we examine the relationships in our study population. The age variable showed a statistically significant correlation between the FEV1/FVC value. FEV1/FVC results varied according to gender differences, these differences were lower than the difference due to weight and age differences. On the other hand, no significant relationship was found between the height variable and the FEV1/FVC value in the analysis we performed on our population. Stronger results can be obtained if the relationships between FEV1/FVC with age, weight, and

sex variables are investigated using a mix regression model in larger populations. We believe it would be more meaningful to consider these relationships when interpreting the FEV1/FVC value.

**Disclosures:** There is no conflict of interest for all authors.

**Acknowledgements:** None.

**Funding:** This research is financially supported by Scientific Research Project (SRP –BAP- FDK-2018-6791), funded by Van YuzuncuYil University, Turkey.

## References

1. Miller, M.R., et al., Standardisation of spirometry. *Eur Respir J.* 2005; 26(2):319-38. <https://doi.org/10.1183/09031936.05.00034805>
2. Ferguson, G.T., et al., Office spirometry for lung health assessment in adults: a consensus statement from the National Lung Health Education Program. *Chest.* 2000; 117(4):1146-1161. <https://doi.org/10.1378/chest.117.4.1146>
3. Minelli, R., Appunti dalle lezioni di fisiologia umana. *La Goliardica Pavese*, 1992: p. Neuroimage.
4. Emmett, M., *Current Clinical Medicine* 2009: Expert Consult Premium Edition by the Cleveland Clinic. Proceedings (Baylor University, Medical Center). 2009; 22(3):291. <https://doi.org/10.1080/08998280.2009.11928536>
5. Johnson, J.D. and W.M. Theurer, A stepwise approach to the interpretation of pulmonary function tests. *Am Fam Physician.* 2014; 89(5):359-66.
6. Vestbo, J., et al., Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am J Respir Crit Care Med.* 2013; 187(4):347-65. <https://doi.org/10.1164/rccm.201204-0596PP>
7. Society, A.T. and M.S.o.t.A.L. Association, Lung function testing: selection of reference values and interpretative strategies. *Am Rev Respir Dis.* 1991; 144(1202):e18. <https://doi.org/10.1164/ajrccm/144.5.1202>
8. Quanjer, P., et al., Standardized lung function testing. *Bull Eur Physiopathol Respir.* 1983; 19(5):1-95.
9. Pearson, M., et al., BTS guidelines for the management of chronic obstructive pulmonary disease-Foreword. *Thorax.* 1997; 52:1-28. [https://doi.org/10.1136/thx.52.suppl\\_5.1](https://doi.org/10.1136/thx.52.suppl_5.1)
10. Stanojevic, S., et al., Reference ranges for spirometry across all ages: a new approach. *American journal of respiratory and critical care medicine.* 2008; 177(3):253-260. <https://doi.org/10.1164/rccm.200708-1248OC>
11. Nunn, A. J., & Gregg, I. New regression equations for predicting peak expiratory flow in adults. *BMJ (Clinical research ed.)*. 1989; 298(6680):1068–1070. <https://doi.org/10.1136/bmj.298.6680.1068>
12. Swanney, M.P., et al., Using the lower limit of normal for the FEV1/FVC ratio reduces the misclassification of airway obstruction. *Thorax.* 2008; 63(12):1046-1051. <http://dx.doi.org/10.1136/thx.2008.098483>
13. Falaschetti, E., et al., Prediction equations for normal and low lung function from the Health Survey for England. *Eur Respir J.* 2004; 23(3):456-63. <https://doi.org/10.1183/09031936.04.00055204>
14. Culver, B.H., Interpretation of spirometry: we can do better than the GOLD standard. *Respiratory Care.* 2006;
15. Hardie, J.A., et al., Risk of over-diagnosis of COPD in asymptomatic elderly never-smokers. *Eur Respir J.* 2002; 20(5):1117-22. <https://doi.org/10.1183/09031936.02.00023202>
16. Wang, P. and M.L. Puterman, Mixed logistic regression models. *Journal of Agricultural, Biological, and Environmental Statistics.* 1998; 175-200. <https://doi.org/10.2307/1400650>
17. Yesilova, A., et al., Locational classification of walnut (*Juglans Regia* L.) genotypes collected from Lake Van basin by using mixture modeling. *African Journal of Agricultural Research.* 2010; 5(12):1509-1514. <https://doi.org/10.5897/AJAR.9000471>
18. Deng, Y. and A. Li, Structural Health Monitoring for Suspension Bridges. Springer. 2019; <https://doi.org/10.1007/978-981-13-3347-7>
19. Wang, Y. and Q. Liu, Comparison of Akaike information criterion (AIC) and Bayesian information criterion (BIC) in selection of stock-recruitment relationships. *Fisheries Research.* 2006; 77(2):220-225. <https://doi.org/10.1016/j.fishres.2005.08.011>
20. Penny, W.D., Comparing dynamic causal models using AIC, BIC and free energy. *Neuroimage.* 2012; 59(1):319-330. <https://doi.org/10.1016/j.neuroimage.2011.07.039>
21. Xie, W., et al., Improving marginal likelihood estimation for Bayesian phylogenetic model selection. *Syst Biol.* 2011; 60(2):150-60. <https://doi.org/10.1093/sysbio/syq085>
22. McDonald, R.P., An index of goodness-of-fit based on noncentrality. *Journal of classification.* 1989; 6(1): 97-103. <https://doi.org/10.1007/BF01908590>
23. Liddle, A.R., Information criteria for astrophysical model selection. *Monthly Notices of the Royal Astronomical Society: Letters.* 2007; 377(1):L74-L78. <https://doi.org/10.1111/j.1745-3933.2007.00306.x>
24. Bozdogan, H. and S.L. Sclove, Multi-sample cluster analysis using Akaike's information criterion. *Annals of the Institute of Statistical Mathematics.* 1984; 36(1):163-180. <https://doi.org/10.1007/BF02481961>
25. Sclove, S.L., Application of model-selection criteria to some problems in multivariate analysis. *Psychometrika.* 1987; 52(3):333-343. <https://doi.org/10.1007/BF02294360>
26. Schwarz, G., Estimating the dimension of a model. *The annals of statistics.* 1978; 6(2):461-464. <https://doi.org/10.1214/aos/1176344136>
27. Sahebajami, H. and P.S. Gartside, Pulmonary function in obese subjects with a normal FEV1/FVC ratio. *Chest.* 1996; 110(6):1425-9. <https://doi.org/10.1378/chest.110.6.1425>

28. García-Río, F., Soriano, J. B., Miravittles, M., Muñoz, L., Duran-Tauleria, E., Sánchez, G., Sobradillo, V., & Ancochea, J. Impact of obesity on the clinical profile of a population-based sample with chronic obstructive pulmonary disease. *PloS one*. 2014; 9(8):e105220. <https://doi.org/10.1371/journal.pone.0105220>
29. Mihmanlı, A., & Bozkurt, E. Effects of Laparoscopic Sleeve Gastrectomy on Respiratory Functions. *Bariatric Surgical Practice and Patient Care*. 2020. <https://doi.org/10.1089/bari.2020.0059>
30. Köchli, S., et al., Lung function, obesity and physical fitness in young children: The EXAMIN YOUTH study. *Respiratory medicine*. 2019; 159:105813. <https://doi.org/10.1016/j.rmed.2019.105813>
31. Çolak, Y., et al., Overweight and Obesity May Lead to Under-diagnosis of Airflow Limitation: Findings from the Copenhagen City Heart Study. *COPD: Journal of Chronic Obstructive Pulmonary Disease*. 2015; 12(1):5-13. <https://doi.org/10.3109/15412555.2014.933955>
32. Dixon, A.E. and U. Peters, The effect of obesity on lung function. *Expert Rev Respir Med*. 2018; 12(9):755-767. <https://doi.org/10.1080/17476348.2018.1506331>
33. Akinbami, O.J. and X. Liu, Chronic Obstructive Pulmonary Disease Among Adults Aged 18 and Over in the United States. 1998-2009. 2011: Citeseer.
34. Runarsdottir, S.B., et al., Prevalence of airflow obstruction in nonsmoking older individuals using different spirometric criteria: the AGES Reykjavik Study. *COPD: Journal of Chronic Obstructive Pulmonary Disease*. 2013; 10(4):493-499. <https://doi.org/10.3109/15412555.2013.773303>
35. Turner, J.M., J. Mead, and M.E. Wohl, Elasticity of human lungs in relation to age. *J Appl Physiol*. 1968; 25(6):664-71. <https://doi.org/10.1152/jappl.1968.25.6.664>
36. Knudson, R.J., et al., Effect of aging alone on mechanical properties of the normal adult human lung. *J Appl Physiol Respir Environ Exerc Physio*. 1977; 43(6):1054-62. <https://doi.org/10.1152/jappl.1977.43.6.1054>
37. Burrows, B., et al., The course and prognosis of different forms of chronic airways obstruction in a sample from the general population. *N Engl J Med*. 1987; 317(21):1309-14. <https://doi.org/10.1056/nejm198711193172103>
38. McHugh, J., Duong, M., Ma, J., Dales, R. E., Bassim, C. W., & Verschoor, C. P. A comprehensive analysis of factors related to lung function in older adults: Cross-sectional findings from the Canadian Longitudinal Study on Aging. *Respiratory medicine*. 2020; 173:106157. <https://doi.org/10.1016/j.rmed.2020.106157>