



## Preference of Multivariate Analysis over Univariate Analysis in Lung Function Studies

Authors

**Dr Kalavathi Lakshmipathy<sup>1</sup>, Dr Ghansham Sharma<sup>2</sup>, Mrs. Swetha B. Lingutla<sup>3</sup>**

<sup>1</sup>MD (Physiology), Director – Professor, Department of Physiology, ESIC Medical College & Post Graduate Institute of Medical Sciences & Research, Rajajinagar, Bangalore-560010, Karnataka, India.

Landline: +91 080 23125572 (Dean), +91 080 23003604 (Chamber)

Mobile No: (+91) 9448259303, Email: [me\\_kalavathi@yahoo.com](mailto:me_kalavathi@yahoo.com)

<sup>2</sup>PhD (Statistics), Assistant Professor, Statistics, Department of Community Medicine, ESIC Medical College & Post Graduate Institute of Medical Sciences & Research, Rajajinagar, Bangalore-560010, Karnataka, India,

Landline: +91 080 23125572 (Dean), Mobile No: (+91) 9620177676, Email: [drghanshamsharma76@gmail.com](mailto:drghanshamsharma76@gmail.com)

<sup>3</sup>M.Sc. (Medical Physiology), Tutor, R.V.M. Institute of Medical Sciences, Laxmakapally, Muluku Mandalam, Medak District, Telangana State, India, Email: [swetha.bindu2@gmail.com](mailto:swetha.bindu2@gmail.com)

### Abstract

Multiple dependent and independent variables affect lung function tests. Hence univariate analysis is not appropriate. The main objectives of the study were to show the appropriateness of multivariate analysis for these types of studies, to know the exact quantification effects of factors on study variables and to study the interaction effects of factors. Eighty individuals between 29-59 years formed the study group. Using computerized spirometer, 3 study variables - Forced Expiratory Volume in 1<sup>st</sup> sec / Forced Vital Capacity, Forced Expiratory Flow rate 25-75% and Peak Expiratory Flow Rate, were determined and correlated to 3 factors - age, gender and height. Seven Models were formulated by different combinations of factors. Each Model was analysed by Multivariate analysis of Variance (MANOVA) which resulted Wilks' Lamda ( $\lambda$ ), Univariate ANOVA with full Factorial Experiments ( $2^n$ ) and K-matrix with Bonferoni's confidence interval. Geometric Mean was calculated from partial  $\eta^2$  values. This methodology is superior and exact compared to univariate analysis. Overall contribution of age in influencing study variables simultaneously is 67.4%, gender 88.3%, height 63.2%, age-gender 30.9%, age-height 28.2% and gender-height 31.5%. This effect quantification information is not available in literature as their analysis was by univariate analysis. All main effects and second order interaction effects of factors are significantly influencing study variables. It is concluded that multivariate analysis is preferred over univariate analysis in lung function studies.

**Keywords:** Wilks Lamda, MANOVA, K -matrix, partial  $\eta^2$ .

### Introduction

With improvisations in spirometry since 1846 lung function tests have evolved as a part of routine health examinations in public health screening<sup>1</sup>. Univariate analysis of the data are plenty in literature<sup>2-9</sup>. The novelty in this study is

Multivariate Analysis of Variance (MANOVA). A systematic analysis of the data to explore linear relationship of Forced Expiratory Volume in 1<sup>st</sup> sec / Forced Vital Capacity (FEV<sub>1</sub>/ FVC %), Forced Expiratory Flow 25-75% (FEF<sub>25-75%</sub>, litres/sec) and Peak Expiratory Flow Rate (PEFR,

litres/sec) with age, gender and height is discussed.

The main objectives of the study were

- 1) To know the exact quantification effects of factors – age, gender and height on study variables by multivariate analysis.
- 2) To study the interaction effects of factors on study variables.

### Materials and Methods

An analytical cross sectional study was undertaken in the Physiology department of a South Indian medical college during 2009 after approval by the Institutional Ethics Committee. South Indians here are of Dravidian origin living in tropical climate at sea level with rice as their staple food.

The sample size of 80 was calculated by considering the mean FVC for 2 age groups (20-39 & 40-59 years) after referring a multivariate study<sup>10</sup>. The mean FVC for age group 1 and 2 are  $4.40 \pm 0.16$  and  $3.79 \pm 0.14$ . The sample size of 40 per group was decided by taking the higher SD, 0.16 with 0.61 effect size at 5% level of significance assuming two tailed hypothesis in such a manner to maintain a power of 99.99%. Subjects of both genders, falling in similar socio economic cohort apparently, were screened through a questionnaire, after written informed consent was taken. Those with history of smoking, recent febrile illness, history of cardiac, respiratory, neuromuscular diseases, diabetes, and asthma were excluded from the study.

The subjects were assessed using a computerized spirometer medspirer (RMS Helios 401) which has pulmonary calculator system. This system has flow transducer to measure flow which is integrated by computer to volume. The system plots a point to point flow volume curve. The instrument was fully calibrated every day. The subjects reported for the department between 11am - 3pm and were made familiar with the technique. They completed atleast 3 acceptable maneuvers in sitting posture and the best maneuver was taken as standard.

The present study involved 3 factors at 2 levels each – Age ( $F_1$ ), 20-39 years ( $A_1$ ) & 40-59 years ( $A_2$ ); Gender ( $F_2$ ), Male & Female and Height ( $F_3$ ), 141-159 cms ( $H_1$ ) & 160–179cm ( $H_2$ ). Each level had 40 subjects except Height where  $H_1$  and  $H_2$  had 42 and 38 subjects respectively.  $F_1$  and  $F_3$  are continuous covariates on ratio scale and  $F_2$  is on nominal scale. For statistical analysis, these ratio scale covariates were transformed to ordinal scale. Study variables (V) were:  $V_1$ - FEV<sub>1</sub>/ FVC,  $V_2$ - FEF<sub>25-75%</sub> and  $V_3$ - PEFR.

Possible Models with all combinations of factors to study their effects on study variables were formulated as, Model 1- Age; Model 2- Gender; Model 3- Height; Model 4- Age-Gender; Model 5 - Age-Height ; Model 6- Gender-Height and Model 7 - Age- Gender-Height. Each Model was tested for assumption of normality and equal variance by Levine test and Spread Vs Level graph. Distribution differed from Model to Model, as and when a factor was added or deleted. If assumption was not met with original observed values of study variables, it was retested after transformation. Best suited transformations were log and reciprocal. Transformed values were utilized only for the purpose of statistical analysis. Inferences were drawn only on the basis of corresponding original values. This is a well-established and accepted procedure in statistics<sup>11,12</sup>. This was challenging and was done to study the effects of factors  $F_1$ ,  $F_2$  and  $F_3$  on  $V_1$ ,  $V_2$  and  $V_3$  simultaneously (all study variables in single group).

### MANOVA Hypothesis

$$\text{Model 1 } H_0: \mu_{A1} = \mu_{A2}$$

$$\text{Model 2 } H_0: \mu_M = \mu_F$$

$$\text{Model 3 } H_0: \mu_{H1} = \mu_{H2}$$

$$\text{Model 4 } H_0: \mu_{A1M} = \mu_{A2M} = \mu_{A1F} = \mu_{A2F}$$

$$\text{Model 5 } H_0: \mu_{A1H1} = \mu_{A2H1} = \mu_{A1H2} = \mu_{A2H2}$$

$$\text{Model 6 } H_0: \mu_{H1M} = \mu_{H1F} = \mu_{H2M} = \mu_{H2F}$$

$$\text{Model 7 } H_0:$$

$$\mu_{A1MH1} = \mu_{A1FH1} = \mu_{A2MH1} = \mu_{A2FH1} = \mu_{A1FH2} = \mu_{A1MH2} = \mu_{A2FH2} = \mu_{A2MH2}$$

H<sub>1</sub>: At least one pair is significant.

where  $\mu_s$  are the mean column matrices of the study variables V,

This study tests the above mentioned hypothesis by MANOVA. If null hypothesis is rejected, then atleast one factor is influencing a study variable significantly.

Wilks Lamda Value is a multivariate test statistic. It was tested for significance at 5% level. This is central in multivariate analysis and is critically absent in univariate analysis. If p-value <.05, the factor is significantly influencing the variables under study. Wilks  $\lambda$  came out to be significant meaning at least 1 factor is influencing the study variables. Univariate ANOVA with full factorial experiment (2<sup>n</sup>) analysis was done to find out which factor was influencing. Factorial Experiments (2<sup>n</sup>) are carried out through ANOVA where n=3, the number of factors. Hypothesis for V<sub>1</sub> are:

Model 1: For V<sub>1</sub>:- H<sub>0</sub>:  $\mu_{A1} = \mu_{A2}$  (for 2 levels),

H<sub>0</sub>:  $\mu_{A1} = \mu_{A2} = \mu_{A3}$  (if there are 3 levels)

Model 2: For V<sub>1</sub>:- H<sub>0</sub>:  $\mu_M = \mu_F$

Model 3: For V<sub>1</sub>:- H<sub>0</sub>:  $\mu_{H1} = \mu_{H2}$

Model 4: For V<sub>1</sub>:- H<sub>0</sub>:  $\mu_{A1M} = \mu_{A2M} = \mu_{A1F} = \mu_{A2F}$

Model 5: For V<sub>1</sub>:- H<sub>0</sub>:  $\mu_{A1H1} = \mu_{A2H2} = \mu_{A1H2} = \mu_{A2H1}$

Model 6: For V<sub>1</sub>:- H<sub>0</sub>:  $\mu_{MH1} = \mu_{MH2} = \mu_{FH1} = \mu_{FH2}$

Model 7: For V<sub>1</sub>:- H<sub>0</sub>:  $\mu_{A1MH1} = \mu_{A2MH1} = \mu_{A1FH1} = \mu_{A2MH2} = \mu_{A1MH2} = \mu_{A2FH2} = \mu_{A2FH1} = \mu_{A1FH2}$

H<sub>1</sub>: At least one pair is significant.

**Table 1:** Descriptive statistics

| Factor                             | Number of subjects (n) | Mean ± SD                 |                               |              |
|------------------------------------|------------------------|---------------------------|-------------------------------|--------------|
|                                    |                        | FEV <sub>1</sub> /FVC (%) | FEF <sub>25-75%</sub> (l/sec) | PEFR (l/sec) |
| Level 1 – A <sub>1</sub> (20-39)   | 40                     | 83.11 ± 2.75              | 3.56 ± 0.75                   | 7.49 ± 1.31  |
| Level 2 – A <sub>2</sub> (40-59)   | 40                     | 78.52 ± 1.91              | 2.92 ± 0.68                   | 7.1 ± 1.44   |
| Total                              | 80                     | 80.81 ± 3.3               | 3.24 ± 0.78                   | 7.3 ± 1.38   |
| Gender                             |                        |                           |                               |              |
| Level 1 – Male                     | 40                     | 82.64 ± 3.21              | 3.90 ± 0.44                   | 8.56 ± 0.59  |
| Level 2 – Female                   | 40                     | 79.00 ± 2.21              | 2.56 ± 0.35                   | 6.01 ± 0.44  |
| Total n=80                         | 80                     | 80.82 ± 3.3               | 3.24 ± 0.78                   | 7.29 ± 1.38  |
| Height (in cms)                    |                        |                           |                               |              |
| Level 1 – H <sub>1</sub> (141-159) | 42                     | 80.63 ± 3.54              | 2.87 ± 0.73                   | 6.41 ± 0.99  |
| Level 2 – H <sub>2</sub> (160-179) | 38                     | 81.02 ± 3.04              | 3.64 ± 0.64                   | 8.25 ± 1.08  |
| Total                              | 80                     | 80.82 ± 3.3               | 3.24 ± 0.78                   | 7.29 ± 1.38  |

Similarly we can write hypothesis for V<sub>2</sub> and V<sub>3</sub> in all the Models.

ANOVA result was significant and hence we proceeded further for K matrix to know which pair of level within a factor is significantly influencing the variables independently.

### K Matrix for Contrast Comparison

First it is obtained with transformed values. Subsequently, all the values were transformed to original values for the sake of interpretation. The value obtained is the difference of the means of 2 levels.

Hypothesis For age:

Model 1: For V<sub>1</sub>:- H<sub>0</sub>:  $\mu_{A1} = \mu_{A2}$  (for 2 levels)

H<sub>0</sub>:  $\mu_{A1} = \mu_{A2} = \mu_{A3}$  and  $\mu_{A1} = \mu_{A3}$  and  $\mu_{A2} = \mu_{A3}$  (for 3 levels)

H<sub>1</sub>: Atleast one pair is significant

Similarly for V<sub>2</sub> and V<sub>3</sub> in all other Models.

Bonferoni's method for constructing simultaneous confidence interval adjusts according to the number of comparisons. In this procedure, type I error does not exceed beyond the specified level of 5%. Quantification of overall contribution of each factor on the study variables was obtained by calculating geometric mean from their parital  $\eta^2$  values in participating Models.

### Results and Discussion

Table 1 shows the descriptive statistics of the study group. Age >40 years and Height < 160cm is negatively correlated. Males have better values than females.

Table 2 shows MANOVA & full Factorial ANOVA analysis. Wilks  $\lambda$  is statistically significant for each factor in all the Models except in Model 7 for age-gender and age-gender-height

factor combinations. Full Factorial ANOVA analysis gives p-values for each factor with each study variable in each Model. This in fact exactly supports the observations of MANOVA.

**Table 2:** Manova & factorial anova

| Models  | factors    | wilks $\lambda$ | mano va f value | Mano va p value | part ial $\eta^2$ | ANOVA F VALUE         |                       |        | ANOVA P VALUE         |                       |             |
|---------|------------|-----------------|-----------------|-----------------|-------------------|-----------------------|-----------------------|--------|-----------------------|-----------------------|-------------|
|         |            |                 |                 |                 |                   | FEV <sub>1</sub> /FVC | FEF <sub>25-75%</sub> | PEFR   | FEV <sub>1</sub> /FVC | FEF <sub>25-75%</sub> | PEFR        |
| Model 1 | Age (A)    | .386            | 40.37           | .000            | .614              | 75.33                 | 16.45                 | 1.63   | .000                  | .000                  | .206        |
| Model 2 | Gender (G) | .137            | 159             | .000            | .863              | 34.80                 | 230.82                | 478.74 | .000                  | .000                  | .000        |
| Model 3 | Height (H) | .366            | 43.81           | .000            | .634              | .286                  | 24.46                 | 62.89  | .594                  | .000                  | .000        |
| Model 4 | A          | .217            | 89.25           | .000            | .783              | 211.08                | 188.60                | 17.92  | .000                  | .000                  | .000        |
|         | G          | .079            | 289.5           | .000            | .921              | 132.56                | 798.41                | 637.74 | .000                  | .000                  | .000        |
|         | A-G        | .691            | 11.05           | .000            | .309              | 9.99                  | 3.08                  | 7.26   | .002                  | .084                  | .009        |
| Model 5 | Age        | .324            | 51.44           | .000            | .676              | 78.35                 | 31.79                 | 6.09   | .000                  | .000                  | .016        |
|         | H          | .295            | 58.84           | .000            | .705              | 1.66                  | 47.46                 | 75.22  | .201                  | .000                  | .000        |
|         | A-H        | .718            | 9.69            | .000            | .282              | 1.92                  | 12.06                 | 11.63  | .170                  | .001                  | .001        |
| Model 6 | G          | .133            | 160.3           | .000            | .867              | 44.97                 | 161.62                | 429.60 | .000                  | .000                  | .000        |
|         | H          | .360            | 43.80           | .000            | .640              | 7.05                  | 2.83                  | 49.38  | .010                  | .097                  | .000        |
|         | G-H        | .636            | 14.11           | .000            | .364              | 7.60                  | 5.94                  | 1.34   | .007                  | .017                  | .251        |
| Model 7 | A          | .364            | 40.78           | .000            | .636              | 112.42                | 106.29                | 27.15  | .000                  | .000                  | .000        |
|         | G          | .118            | 1.75            | .000            | .882              | 134.01                | 445.45                | 481.89 | .000                  | .000                  | .000        |
|         | H          | .441            | 29.53           | .000            | .552              | 16.77                 | 5.82                  | 58.62  | .000                  | .018                  | .000        |
|         | A-G        | .988            | 0.29            | <b>.830</b>     | .012              | 0.27                  | 0.54                  | 0.10   | <b>.604</b>           | <b>.464</b>           | <b>.751</b> |
|         | A-H        | .887            | 2.96            | .038            | .113              | 2.30                  | 3.76                  | 0.18   | <b>.133</b>           | <b>.056</b>           | <b>.672</b> |
|         | G-H        | .728            | 8.74            | .000            | .272              | 0.29                  | 0.39                  | 10.68  | <b>.595</b>           | <b>.536</b>           | .002        |
|         | A-G-H      | .969            | 0.74            | <b>.530</b>     | .031              | 1.10                  | 1.18                  | 0.44   | <b>.298</b>           | <b>.280</b>           | <b>.835</b> |

Table 3 shows K matrix –contrast comparison for main effects in original values at 95% confidence interval. Most of the p-values <.001, which are highly significant for all the factors on all the variables barring, age on PEFR (Model 1) , height on FEV<sub>1</sub>/FVC (Model 3 & 5) and height on FEF<sub>25-75%</sub> (Model 6). It can also be observed the

confidence intervals are narrow meaning the estimates are authentic. The last row shows the median effect of a factor on a study variable which is the final quantification of age, gender and height on V<sub>1</sub>, V<sub>2</sub> & V<sub>3</sub> independently and all are highly significant ( p-value<.001).

**Table 3:** K matrix - Contrast comparison

| Model            | Age<br>Difference of means, p value<br>(Confidence Interval) |                                   |  | Gender<br>Difference of means, p value<br>(Confidence Interval) |                                  |                                  | Height<br>Difference of means, p value<br>(Confidence Interval) |   |                                   |
|------------------|--|-----------------------------------|--|---|----------------------------------|----------------------------------|---|---|-----------------------------------|
|                  | FEV <sub>1</sub> /<br>FVC                                    | FEF <sub>25-<br/>75%</sub>        | PEFR                                     | FEV <sub>1</sub> /<br>FVC                                       | FEF <sub>25-<br/>75%</sub>       | PEFR                             | FEV <sub>1</sub> /<br>FVC                                       | FEF <sub>25-<br/>75%</sub>              | PEFR                              |
| 1                | 4.592,<br>.000<br>(3.54,<br>5.64)                            | 0.649,<br>.000<br>(0.33,<br>0.97) | 0.393,<br><b>.206</b><br>(0.22,<br>1.01) | -   | -                                | -                                | -   | -                                       | -                                 |
| 2                | -  | -                                 | -  | 3.64,<br>.000<br>(2.41,<br>4.87)                                | 1.34,<br>.000<br>(1.17,<br>1.52) | 2.55,<br>.000<br>(2.32,<br>2.78) | -   | -                                       | -                                 |
| 3                | -  | -                                 | -  | -   | -                                | -                                | 0.397,<br><b>.594</b><br>(-<br>1.08,<br>1.87)                   | 0.76,<br>.000<br>(0.45,<br>1.07)        | 1.838,<br>.000<br>(1.38,<br>2.30) |
| 4                | 4.592,<br>.000<br>(3.96,<br>5.22)                            | 1.23,<br>.000<br>(1.19,<br>1.27)  | 1.06,<br>.000<br>(1.03,<br>1.09)         | 3.639,<br>.000<br>(3.01,<br>4.30)                               | 1.53,<br>.000<br>(1.48,<br>1.57) | 1.42,<br>.000<br>(1.38,<br>1.46) | -   | -                                       | -                                 |
| 5                | 6.53,<br>.000<br>(6.53,<br>6.53)                             | 0.74,<br>.000<br>(0.48,<br>1.01)  | 0.53,<br>.016<br>(.11,<br>1.01)          | -   | -                                | -                                | 0,<br><b>.201</b><br>(0,<br>0)                                  | 0.91,<br>.000,<br>(.65,<br>1.17)        | 1.90,<br>.000,<br>(1.48,<br>2.33) |
| 6                | -  | -                                 | -  | 1.05,<br>.000<br>(1.04,<br>1.07)                                | 1.48,<br>.000<br>(1.40,<br>1.58) | 2.18,<br>.000<br>(1.97,<br>2.39) | 1.02,<br>.01<br>(1.04,<br>1.01)                                 | 1.05,<br><b>.097</b><br>(1.01,<br>1.12) | 0.74,<br>.000<br>(.53,<br>.95)    |
| 7                | 4.26,<br>.000<br>(3.46,<br>5.07)                             | 0.64,<br>.000<br>(0.52,<br>0.76)  | 0.521,<br>.000<br>(.32,<br>.72)          | 4.66,<br>.000<br>(3.85,<br>5.46)                                | 1.31,<br>.000<br>(1.19,<br>1.43) | 2.20,<br>.000<br>(1.20,<br>2.39) | 1.65,<br>.000<br>(0.85,<br>2.45)                                | 0.15,<br>.018<br>(0.03,<br>0.27)        | 0.76,<br>.000<br>(.57,<br>.96)    |
| Median<br>Effect | 4.59,<br>.000<br>(3.46,<br>6.53)                             | 0.66,<br>.000<br>(0.33,<br>1.01)  | 0.48,<br>.000<br>(0.11,<br>1.01)         | 3.80,<br>.000<br>(2.41,<br>5.53)                                | 1.33,<br>.000<br>(1.02,<br>1.57) | 2.19,<br>.000<br>(1.2,<br>2.78)  | 1.34,<br>.000<br>(0.85,<br>2.45)                                | 0.76,<br>.000<br>(.03,<br>1.17)         | 1.30,<br>.000<br>(0.53,<br>2.33)  |

Table 4 shows overall contribution of factor/factor combination on all the study variables simultaneously by calculating geometric mean from partial  $\eta^2$  values. Overall contribution of age influencing the study variables simultaneously

is 67.4%, gender 88.3%, height 63.2%, age-gender 30.9%, age-height 28.2% and gender-height 31.5%. It can be noted that in higher order interactions, % contribution decreases.

**Table 4:** Overall contribution of the factor/factor combination

| Model          | Partial $\eta^2$ Values |        |        |            |              |                 |
|----------------|-------------------------|--------|--------|------------|--------------|-----------------|
|                | Age                     | Gender | Height | Age-gender | Age – height | Gender - height |
| 1              | .614                    | -      | -      | -          | -            | -               |
| 2              | -                       | .863   | -      | -          | -            | -               |
| 3              | -                       | -      | .634   | -          | -            | -               |
| 4              | .783                    | .921   | -      | .309       | -            | -               |
| 5              | .676                    | -      | .705   | -          | .282         | -               |
| 6              | -                       | .867   | .640   | -          | -            | .364            |
| 7              | .636                    | .882   | .559   | -          | -            | .272            |
| Geometric Mean | .674                    | .883   | .632   | .309       | .282         | .315            |

Most authors advocate separate regional prediction equations. These prediction equations cannot be used for a long time because of cohort effect<sup>12-14</sup>.

In our study, normality could not be achieved with any type of transformation in any Model, when weight factor was added, giving a clue that weight is not affecting the study variables much<sup>15</sup>. Hence weight as 4<sup>th</sup> factor was removed from our study. Each Model satisfied Levine test and/or Spread Vs Level graph and that normality simply cannot be presumed. The power of the present study is 99.99%. It is the probability to accept the null hypothesis when actually it is true. It controls type II error (beta error). MANOVA keeps Type I error (alpha error, Null rejected when it is true) at  $\alpha\%$  - 5%, for all hypothesis.

This study is a composite study where 7 Models were constructed and MANOVA hypothesis were formulated keeping type I error at 5% level. It is to be observed that although the no. of hypothesis increased, alpha error always remained at 5%. This is the single most accuracy about MANOVA. In Univariate analysis, each hypothesis is tested at 5% Alpha error and 3 hypothesis each for  $V_1$ ,  $V_2$  &  $V_3$  in each Model will cumulate to 15% error. Bonferoni’s method of estimating confidence interval also maintains alpha error throughout at 5%. A narrow confidence interval indicates a very authentic study.

MANOVA filters unnecessary calculations, by Wilks  $\lambda$  & full Factorial ANOVA. If  $\lambda$  is not significant, further analysis is abandoned concluding that lung function parameters are not

influenced by the factors under study. Exact Quantification of the influence exerted by the factors on the study variables will be obtained using multivariate analysis which is not possible in univariate analysis.

In our study males had better lung functions than females, which was negatively correlated with age and positively correlated with height. Age >40 years and height < 160 cm have negative impact on lung function tests. By magnitude, independently,  $V_1$  was affected most by Age and both  $V_2$  &  $V_3$  by Gender. Overall contribution of factor/factor combination influencing study variables V, simultaneously was - Age -67.4%; Gender- 88.3%; Height- 63.2%; Age-Gender - 30.9%; Age-Height - 28.2% and Gender-Height 31.5%. A factor exerted variable influences in different Models on the same study variables.

Exhaustive Indian studies<sup>15,16</sup> had derived multivariate prediction equations and compared with older equations. Facts like autocorrelation and adjusted  $R^2$  were overlooked. In the Multivariate Analysis study, carried out by Verma et al in 2002<sup>10</sup>, the stratification of age into 7 groups was made, each group having 20-30 subjects, with 6 study variables. It is unlikely that assumption would be satisfied. Good attempt had been done by calculating  $\lambda$  value and Mahalanobi’s Distance to test the significance between any 2 age groups. In reality age alone is not an independent predictor. Japanese studies of Discriminant analysis<sup>17-19</sup> were really credible.

### Conclusion

MANOVA is preferred over ANOVA in lung function studies, as the methodology is appropriate, superior and exact quantification of the influence exerted by a factor is obtained. All the main effects and second order interaction effects of factors Age, Gender and Height are significantly influencing study variables. Their influence decrease for higher order interactions. The factors exert variable influences in different Models on the same study variables. No single Model is reliable.

### References

1. Spriggs E. A., Gandevia B. and Bishop P. J. John Hutchinson (1811-1861) Inventor of the Spirometer. Proceedings of Royal Society of Medicine. 1976; 69(6): 450–450.
2. Dikshit M. B., Raje S. and Agrawal M. J. Lung functions with spirometry: An indian Perspective-1. Peak Expiratory Flow Rates. Indian Journal of Physiology and Pharmacology. 2005a; 49(1): 8–18.
3. Dikshit M. B., Raje S. and Agrawal M. J. Lung functions with spirometry: An indian Perspective-II. Vital Capacity of Indians. Indian Journal of Physiology and Pharmacology. 2005b 49(3): 257–270.
4. Tsukioka K., Makino S., Miyamoto T., Tanabe N. and Akazawa K. Peak Expiratory flow in normal healthy Japanese adults: Comparisons with values reported in other races. Allergology international. 2002; 51: 101–111.
5. Vijayan V. K., Kuppurao K. V., Venkatesan P., Sankaran K. and Prabhakar R. Pulmonary functions in healthy young adult Indians in Madras. Thorax. 1990; 45: 611–615.
6. Udawadia F., Sunavala J. D., Shetye M. and Jain P. K. The maximal flow volume curve in normal subjects in India. Chest. 1986; 89: 852–857.
7. Williams D. E., Miller D. R. and Taylor F. W. Pulmonary function studies in healthy Pakistani adults. Thorax. 1978; 33: 243–249.
8. Dacosta J. L. Pulmonary function studies in healthy chinese adults in Singapore. American Review Respiratory Disease. 1971; 104: 128–31.
9. Jain S. K. and Guptha C. K. Lung function studies in healthy men and women over forty. Indian Journal of Medical Research. 1967; 55: 612–619.
10. Verma S. S., Sharma Y. K. and Arora S. A. Multivariate study of some lung function tests at different age groups in healthy Indian males. Indian Journal of Chest Diseases and Allied Sciences. 2002; 44: 85–90.
11. Fernandez G. C. J. Residual analysis and Data Transformations: Important Tools in Statistical Analysis. Horticultural Science. 1992; 27(4): 297–300.
12. Society, The American Thoracic, O. statement of the American Thoracic Society. (1991). Lung Function Testing: Selection of Reference Values and Interpretative Strategies. American Review Respiratory Disease. 1991; 144: 1202–18.
13. Xu X., Laird N., Dockery D. W., Schouten J. P., Rijcken B. and Weiss S. T. Age, period and cohort effects on pulmonary function in a 24 year longitudinal study. American Journal of Epidemiology. 1995; 141: 554–66.
14. Glindmeyer H. W., Diem J. E., Jones R. N. and Weill H. Noncomparability of longitudinally and cross-sectionally determined annual change in spirometry. American Review of Respiratory Disease. 1982; 125: 544–8.
15. Chhabra S. K., Kumar R., Gupta U., Rahan M. and Dash J. D. Prediction Equations for Spirometry in Adults from Northern India. Indian Journal of Chest

- Disease and Allied Journal Sciences. 2014; 56: 221–229.
16. Dasgupta A., Dhar R., Ghoshal A. G., Kundu S., Mukherjee S., Mukhopadhyay A. and Roychowdhury S. Reference equation for spirometry interpretation for Eastern India. Lung India. 2015; 32(1): 34–39.
  17. Meguro O. T. Discriminant analysis of pulmonary function parameters – mild asthmatics versus moderate asthmatics. Acta Medica Okayama. 1983; 37(1): 51–58.
  18. Meguro T. Selection of effective maximal expiratory parameters to differentiate asthmatic patients from healthy adults by discriminant analysis using all possible selection procedure. Acta Medica Okayama. 1986; 40(4): 215–23.
  19. Meguro O. T. and Tsubota N. (1978). Discriminant analysis of bronchial asthma by linear discriminant function with parameters of flow-volumes: discriminant analysis of bronchial asthma in young male non-smokers. Acta Medica Okayama. 1978; 32(5): 355–361.