5-14-1991

Cystic Fibrosis: Cluster Analysis of Microbiology and Pulmonary Function

Lori Lee M. Larson '91
Illinois Wesleyan University

Cystic Fibrosis: Cluster Analysis of Microbiology and Pulmonary Function

Lori Lee M. Larson
Research Honors
Departments of Biology & Computer Science
Research Advisor, Dr. Lisa Brown
May 14, 1991

Abstract

In an attempt to find a relationship between pulmonary infection and pulmonary function in cystic fibrosis (CF) patients, microbiology data and pulmonary function test (PFT) data for clinic patients' visits were obtained from the University of Minnesota Relational Database. The two files were merged, totaling 12,193 cases, and then analyzed using a quick cluster subroutine of SPSSX on the University of Illinois IBM Mainframe System. QUICK CLUSTER analysis showed a relationship between the virulence of the microorganisms, the amount of growth of the microorganisms, and the pulmonary function test scores of CF patients

Background

Cystic fibrosis (CF) is an autosomal recessive disease. The loss of a single specific trinucleotide codon accounts for about 70% of CF cases. The gene involved codes for an ion channel regulator protein called the Cystic Fibrosis Transmembrane Regulator protein (CFTR). The production of defective CFTR proteins results in abnormal regulation of chloride ion channels which leads to abnormally thick mucus secretions (1). Viscid mucus secretions obstruct the lungs and provide a suitable environment for infection by microorganisms. Infection, in turn, causes mucus hypersecretion. Recurrent pulmonary infection results in destruction of lung parenchyma (2).

The original aim of this study was to determine if there were correlations between clinical microbial data and home monitoring data. Home monitoring allows the physician to follow the patient's health status while the patient is at home instead of in the hospital or clinic (3). Sometimes, changes in the CF patient's health occur slowly and may go unrecognized by the patient or parent/support team resulting in an increased morbidity and/or mortality (4). It was hoped that the correlations would provide earlier detection of negative trends and confirm stability or improving health between clinic visits. With earlier detection of deteriorating health status, treatment becomes more effective. This is especially important today in light of recent developments in cystic fibrosis research, and the possibility of gene therapy in the near future.

Ninety percent of all CF deaths are due to respiratory failure (5). The three bacterial organisms that chronically colonize the

airways of cystic fibrosis patients are *Staphylococcus aureus*, *Hemophilus influenzae*, and *Pseudomonas aeruginosa* (6). Other microorganisms, for example, fungi such as *Aspergillus* and *Candida albicans*, also are common in CF sputum (7). Clinical symptoms of infection include increasingly productive cough with yellow or green sputum, wheeze, weight loss, and fever (5). These symptoms were included in home monitoring.

For the purposes of the original proposal, pulmonary function tests (PFTs) would have been used as a measure of wellness. If the research had shown a relationship between the microbial data and the PFT data and the same relationship between the home data and the PFT data, then, perhaps, the home data could have been used as a predictor of patient status between clinic visits. Unfortunately, the time remaining to complete the research project was not sufficient to be able to combine the computer files containing the microbe data, the home monitoring data, and the PFT data. Therefore, it was decided to investigate the relationships between the microbiology and the patient's health status, as indicated by PFT scores, by using QUICK CLUSTER analysis.

Methods

One hundred and seventy three CF patients receiving treatment from the Cystic Fibrosis Center at the University of Minnesota volunteered to record daily symptoms and measurements in a home health diary. These data were sent to the CF Center weekly and were entered into the (IDM/IBM) relational database system. The patient's status according to the attending physician and other data, such as PFT scores and microorganisms grown in culture, were also entered for each clinic visit.

A visit to the University of Minnesota in October 1990 led to a meeting with several of the University Staff, Dr. Warren Warwick, Dr. Stan Finkelstein and Cathy Wielinski, and a discussion concerning the data that would be retrieved. In January 1991, the requested data from the entire time span of the relational database (1982-1986) were received. These included microbiology and pulmonary function information from each clinic visit, and home data from two weeks before and two weeks after each clinic visit. The home data were collected for this time span because CF patients were more apt to complete the home diaries at these times. In addition, the home data from this time period are more likely to indicate a change in patient health status because treatment is usually altered by the physician at the clinic visit.

SPSSX was used for statistical analysis of the data. The format of the microbiology data was altered in order to be compatible with SPSSX requirements. That is, it had to be changed such that each variable had the same number of characters. SPSSX on the University of Illinois IBM Mainframe System was used to

analyze the data because Illinois Wesleyan's PC version could not handle the large amount of data collected.

The microbiology file, MICRO DAT A, contained the following data.

1. microbe name
2. microbe strain
3. source of culture
   a) tr = trachea
   b) sp = sputum
   c) th = throat
   d) ur = urine
   e) bl = blood
   f) br = bronchial washing
   g) ot = other
4. microbe growth
   a) l = light
   b) m = moderate
   c) h = heavy

Because the SPSSX procedures used require all variables to be numeric, the growth of the microbes, the source of culture , and the microorganisms themselves had to be assigned numeric values. In addition to being assigned a numeric value, the microorganisms were ranked according to their degree of virulence by Dr. Warwick (11). The microbes ranged from 1 being the mildest to 8 being the most severe with regard to degree of resulting pulmonary infection.

1. growth
   a) light = 1
   b) moderate = 2
   c) heavy = 3
2. source
   a) tr = 1
   b) sp = 2
   c) th = 3
   d) ur = 4
   e) bl = 5

f) br = 6
g) ot = 7

3. microbes (name/numerical value/virulence rankage)
Acremonium species / 1 /1
Acinetobacter calcoaceticus v. lwoffii / 2 /1
Acinetobacter calcoaceticus v. anitratus / 3 / 1
Acinetobacter species / 4 / 1
Acinetobacter calcoaceticus / 5 / 1
Achromobacter xylosoxidans / 6 / 1
Achromobacter species / 7 / 1
Aeromonas hydrophilia / 8 / 1
Agrobacterium radiobacter / 9 / 1
Alternaria species / 10 / 1
Alcaligenes species / 11 / 1
Alcaligenes odorans / 12 / 1
Alcaligenes denitrificans / 13 / 1
Aspergillus fumigatus / 14 / 1
Aspergillus versicolor / 15 / 1
Aspergillus terreus / 16 / 1
Aspergillus flavus / 17 / 1
Aspergillus clavatus / 18 / 1
Aspergillus niger / 19 / 1
Bacillus species (gram -) (non-fermenter) / 20 / 1
Branhamella Catarrhalis / 21 / 1
Bacillus species-not *anthracis* / 22 / 1
Bacillus species (gram -) / 23 / 1
Bacillus species / 24 / 1
Candida parapsilosis / 25 / 1
Candida albicans / 26 / 1
Candida tropicalis / 27 / 1
Candida lusitaniae / 28 / 1
Candida rugosa / 29 / 1
Cladosporium species / 30 / 1
Citrobacter amalonaticus / 31 / 1
Citrobacter freundii / 32 / 1
Citrobacter diversus / 33 / 1
CDC Group VE-2 (non-fermenter) / 34 / 1
Enterobacter taylorae / 35 / 1
Enterobacter agglomerans / 36 / 1
Enterobacter aurogenes / 39 / 1
Enterobacter sakazakii / 40 / 1
Fungus (filamentous) / 41 / 1
Fusarium species / 42 / 1

Flavobacterium species / 43 / 1
Hafnia alvei / 45 / 1
Klebsiella oxytoca / 46 / 1
Klebsiella pneumonia / 47 / 1
Klebsiella ozaenae / 48 / 1
Moraxella phenylpyruvica / 49 / 1
Mucor species / 50 / 1
Moraxella species / 51 / 1
Morganella morganii / 52 / 1
Orzihabitans (pseudo) / 53 / 1
Petriellidium boydii /71 / 1
Penicillium species / 72 / 1
Providencia rettgeri / 73 / 1
Providencia stuartii / 74 / 1
Proteus vulgaris / 75 / 1
Proteus mirabilis / 76 / 1
Pasteurella multocida / 77 / 1
Rod (Lactose-) (non-fermenter) / 78 / 1
Serratia marcescens / 79 / 1
Serratia liquefacians / 80 / 1
Saprophyte non-sporulating / 81 / 1
Scapulariopsis species / 82 / 1
Streptococcus beta (group A) / 83 / 1
Streptococcus beta (non-group A) / 84 / 1
Streptococcus species / 85 / 1
Streptococcus pneumonia (Diplococcus) / 86 / 1
Streptococcus beta hemolytic (group B) / 87 / 1
Streptococcus beta hemolytic (group C) / 88 / 1
Streptococcus beta hemolytic (group G) / 89 / 1
Streptocuccus beta hemolytic / 90 / 1
Saccharomyces cerevisiae / 94 / 1
Trichosporon beigelii / 95 / 1
Trichosporon species / 96 / 1
Trichoderma species / 97 / 1
Unspecified Lactose non-fermenter / 98 / 1
Unspecified Lactose fermenter / 99 / 1
Escherichia coli / 37 / 2
Escherichia coli (mucoid) / 38 / 2
Staphylococcus aureus / 91 / 3
Staphylococcus species (Coag-) / 92 / 3
Staphylococcus epidermidis / 93 / 3
Hemophilus influenzae / 44 / 4
Pseudomonas species / 54 / 5

Pseudomonas species (mucoid) / 55 / 5
Pseudomonas species (brown pig) / 56 / 5
Pseudomonas species (fluorescen) / 57 / 5
Pseudomonas thomasii / 61 / 5
Pseudomonas stutzeri / 62 / 5
Pseudomonas alcaligenes / 63 / 5
Pseudomonas putida / 64 / 5
Pseudomonas pickettii / 65 / 5
Pseudomonas fluorescense / 66 / 5
Pseudomonas fluorescense (mucoid) / 67 / 5
Pseudomonas paucimoblis / 68 / 5
Pseudomonas luteola / 69 / 5
Pseudomonas medoncina / 70 / 5
Pseudomonas aeruginosa / 59 / 6
Pseudomonas aeruginosa (mucoid) / 60 / 7
Pseudomonas cepacia / 58 / 8

The pulmonary function file, PFT OUT B, contained the following data.

FEV = (FEF1/FVC)

where FEF1 = forced expiratory flow rate in 1 sec.
FVC = forced vital capacity [the quantity of air expired forcefully after inspiring maximally (9)]
FEV = forced expiratory volume (FEF1 as percent of total FVC)

FVC = (FVC/PFVC)

where FVC = percent forced vital capacity of predicted

FEF1 = (FEF1/PFEF1)

where FEF1 = percent forced expiratory flow in 1 sec. of predicted

MICRO DAT A and PFT OUT B were sorted by patient identification number and clinical date and were merged into one file, CLINIC OUT B. There was a total of 12,193 cases. Because

there was such a large amount of data, QUICK CLUSTER analysis was used to find relationships between pulmonary infection and pulmonary function. First, three clusters for each of the variables were determined. Next, the center, a vector, for each of the three clusters for each variable was found. Finally, the variables were regrouped according to the cluster centers.

Results and Discussion

Appendices A, B, C, and D show the results of QUICK CLUSTER analysis. Initial Cluster Centers are the original three cluster groups for each variable. Classification Cluster Centers are the centers of the three cluster groups for each variable. Final Cluster Centers are the final three cluster groups for each variable after the variables are regrouped according to the cluster centers. IMPORTM is the average value of the codes for virulence for each cluster. CODES is the average value of the codes for the culture source for each cluster. CODEG is the average value of the codes for the growth of the microorganisms for each cluster. The average values for FEV, FVC, and FEV1 for each cluster are also shown.

QUICK CLUSTER analysis indicated relationships between pulmonary infection and pulmonary function. Appendix A shows the results when IMPORTM, CODES, STRAIN, CODEG, FEV, FVC, and FEV1 were clustered together. As the severity of the microorganisms increased, there was a decrease in pulmonary function indicated by decreasing PFT values. In addition, as the growth of the microbes became heavier, the PFT values decreased. Table 1 is a summary of Appendix A.

Table 1 Relationship Between Virulence (IMPORTM), Growth (CODEG), & Pulmonary Function (FEV, FVC, FEV1)

Cluster	IMPORTM	CODEG	FEV	FVC	FEV1
3	2.7054	1.2900	82.4747	111.3475	100.3203
1	3.5916	1.7556	65.6977	89.1721	65.4064
2	4.0883	2.0497	50.3027	52.0539	29.4637

[Cluster 3 = mild pulmonary involvement]
[Cluster 1 = moderate pulmonary involvemen]
[Cluster 2 = severe pulmonary involvement]

CODES for each cluster in Appendix A has a value close to 2 because most cultures were obtained from sputum. This is because pulmonary infection in cystic fibrosis patients does not appear to spread systemically (2).

Appendix B shows the results when IMPORTM, CODEG, FEV, FVC, and FEV1 were clustered together. A similar relationship as in Appendix A existed between the microbes and the PFTs and between the growth of the microbes and the PFTs. Table 2 is a summary of Appendix B.

Table 2 Relationship Between Virulence (IMPORTM), Growth (CODEG), & Pulmonary Function (FEV, FVC, FEV1)

Cluster	IMPORTM	CODEG	FEV	FVC	FEV1
1	2.3766	1.7042	83.2243	114.4795	104.0333
3	2.9295	2.1171	69.5367	91.0235	70.1553
2	3.3933	2.6035	45.5300	52.8268	26.7922

[Cluster 1 = mild pulmonary involvement]
[Cluster 3 = moderate pulmonary involvement]
[Cluster 2 = severe pulmonary involvement]

Appendix C shows the results when IMPORTM, FEV, FVC, and FEV1 were clustered together. Note that increased microbe severity does correspond to decreased PFT values. Table 3 is a summary of Appendix C.

Table 3 Relationship Between Virulence (IMPORTM) & Pulmonary Function (FEV,FVC, FEV1)

Cluster	IMPORTM	FEV	FVC	FEV1
2	2.4158	82.7902	113.4465	102.6389
3	2.9335	68.9998	90.1445	68.8836
1	3.4043	45.4341	53.0171	26.8405

[Cluster 2 = mild pulmonary involvement]
[Cluster 3 = moderate pulmonary involvement]
[Cluster 1 = severe pulmonary involvement]

Although, the differences in severity do not seem to be as large as one might expect, there are number of reasons why the differences are slight. First, though severity was ranked from 1 through 8, the actual number of mild cases was much larger than the number of severe cases. This is evident in Appendix E which shows the frequency distribution of the microorganisms after they were ranked according to virulence. Note too that Appendix E shows that the microbe ranked with a severity of 4, *Hemophilus influenzae*, has no cases because of a misspelling in the coding section of the program. The results were affected since this microorganism is very common in CF sputum. In addition, increasing the number of cluster groups might lead to fewer misclassifications of the variables and thus, improve results. Lastly, there is the possibility that one cannot find a strong relationship between PFT values and microbe virulence alone.

A question may be raised as to whether or not the ranking in virulence of the microorganisms might be based on past pulmonary function test results, thus influencing the results of QUICK CLUSTER analysis. However, if this were the case, one would expect a much larger difference between the clusters indicating the relationship between IMPORTM and PFT values. Instead, the rankings were made primarily on the ability of the microorganism to invade the body, resist treatment, and produce varying degrees of harmful effects on the host.

Staphylococcus aureus is usually the first organism to cause pulmonary infection in cystic fibrosis patients. The bacteria injures the pulmonary tissue by producing toxins, such as hyaluronidase and necrotoxins, and by eliciting the host inflammatory response (10).

Furthermore, mucus hypersecretion and bronchoconstriction are stimulated by the presence of bacterial antigens (2). *Staphylococcus aureus* is found in patients with mild pulmonary disease. *Hemophilus influenzae* and *Pseudomonas aeruginosa* appear after the initial infection has occurred and are associated with moderate to severe pulmonary disease. The conversion of *Pseudomonas* to a mucoid form is a unique characteristic of cystic fibrosis (10). It is found in 90 percent of all CF patients because once infection has occurred , the bacteria are rarely eliminated (6). *Pseudomonas aeruginosa* produces a dense capsule (alginate) which protects the bacteria from the host's immune system and antibiotics (10). Mucoidy also helps the bacteria to adhere to the respiratory tract and seems to provide a selective advantage for colonization and obstruction of the small airways (5). In addition, several enzymes produced by this bacteria can injure the pulmonary tissue. Colonization by *Pseudomonas cepacia* is a poor prognostic sign because this bacteria is resistant to many common antipseudomonal antibiotics (7).

PFTs are used to evaluate treatment and progression of pulmonary involvement (11). A decrease in FEV1 is an indicator of deterioration of pulmonary of function. In general, it is predicted that the CF patient will be alive in four years if the FEV1 is greater than 60% of the predicted value. If the FEV1 is less than 35% of predicted, however, the patient has only a 60% chance of being alive in the next four years (7). Figure 1, taken from Simmons Current Pulmonology (2), relates well with the QUICK CLUSTER results. The values given for FEV1 in Figure 1 are similar to the FEV1 values for

mild, moderate, and severe lung involvement given in Tables 1, 2, 3, and 4. The much lower value for severe lung involvement determined by the QUICK CLUSTER analysis, compared to that given in Figure 1, may be due to the missing *Hemophilus influenzae* data.

Figure 1 "Progressive Pulmonary Function Abnormalitites in Cystic Fibrosis" (2)

		Lung Involvement			
		Normal	*Mild*	*Moderate*	*Severe*
FEV_1		Normal	Normal	70% Pred.	40% Pred.
RV/TLC		25%	35%	55%	70%
FEF_{25-75}		Normal	60% Pred.	40% Pred.	20% Pred.
MEFV Curve					
Broncho-reactivity	Exercise	Neg.	Greater rise during exercise	Variably Positive	Positive
	Histamine	Neg.	Neg.	Variably Positive	Positive
P_aO_2		N	↓	↓↓	↓↓↓
P_aCO_2		N	N	N	↑
Age (yrs)		5	10	15	20

FIG 8.
Representative progress of deterioration of pulmonary function in a subject. FEV_1 = forced expiratory volume at 1 second; *RV/TLC* = residual volume/total lung capacity; FEV_{25-75} = forced expiratory flow over the middle half of the total volume exhaled; *MEFV* = maximal expiratory flow–volume curve; *Pred.* = predicted; *N* = normal. (From Taussig LM, Landau LI, Marks MI: Respiratory system, in Taussig LM (ed): *Cystic Fibrosis*. New York, Thieme-Stratton Inc, 1984, pp 115–174. Reproduced by permission.)

Appendix D shows the results when CODEG, FEV, FVC, and FEV1 are clustered together. Note that an increase in microbe growth does relate to a decrease in PFT values. Table 4 is a summary of Appendix D.

Table 4 Relationship Between Growth (CODEG) & Pulmonary Function (FEV, FVC, FEV1)

Cluster	CODEG	FEV	FVC	FEV1
1	1.6930	82.7923	113.4438	102.6395
3	2.1535	68.9982	90.1464	68.8832
2	2.6029	45.4341	53.0171	26.8405

[Cluster 1 = mild pulmonary involvement]
[Cluster 3 = moderate pulmonary involvement]
[Cluster 2 = severe pulmonary involvement]

Originally, CORRELATIONS analysis was attempted (see Appendix F). However, little correlation was actually found between IMPORTM and PFT values or between CODEG and PFT values (Table 5). Still, the correlation coefficients do show a negative correlation which was expected.

Table 5 Correlation Coefficients for Virulence (IMPORTM), Growth (CODEG), & Pulmonary Function (FEV, FVC, FEV1)

	FEV	FVC	FEV1
IMPORTM	-.1880	-.1277	-.1661
CODEG	-.2853	-.1984	-.2562

[correlation coefficient 0 = no correlation]
[correlation coefficient 1 = perfect positive correlation]
[correlation coefficient -1 = perfect negative correlation]

A possible reason for these results is that the predictor variables, IMPORTM and CODEG, have to few categories while the predicted variables, FEV, FVC, FEV1, have extremely wide ranges. Thus, little correlation was shown.

Conclusion

QUICK CLUSTER analysis of cystic fibrosis clinical microbial data and pulmonary function tests indicate a relationship between the virulence of the microorganisms and/or the amount of growth of the microorganisms and pulmonary function. Perhaps these results, in conjunction with the home monitoring data, could be used to predict when more intensive treatment should be used to prevent further deterioration of pulmonary function. The cluster analyses are only exploratory; further analysis, including program error corrections, is required for more definite conclusions.

References

1. Marx, J. L. "The Cystic Fibrosis Gene is Found," Science. Vol. 245, 1 September, 1989, pp. 923-925.

2. Simmons, D. H., ed. Current Pulmonology. Vol. X, Chicago: Year Book Medical Publishers, Inc., 1989, pp. 377-416.

3. Brown-Ewing, L. J. The Development of an Early Detection and Intervention Criterion for Cystic Fibrosis Patient Clinical Status Based on Home Monitoring Data. University of Minnesota Doctor's Thesis, June 8, 1986, pp. 1-11, 2-1.

4. Brown-Ewing, L. J., S. M. Finkelstein, J. R. Budd, S. Rich, S. J. Kujawa, C. L. Wielinski, and W. J. Warwick. "A Rule for the Early Detection of Chronic Changes in Cystic Fibrosis Patient Status," Journal of Clinical Epidemiology. Vol. 41, No. 9, 1988, pp. 915-922.

5. Fick, R. B., Jr., ed. "Inflammatory Disorders of the Airways," Clinics in Chest Medicine. Vol. IX, No. 4, December 1988, pp. 679-687.

6. Braunwald, E., K. J. Isselbacher, R. G. Petersdorf, J. D. Wilson, J. B. Martin, and A. S. Fauci, eds. Harrison's Principles of Internal Medicine Vol. II. 11th ed., New York: McGraw-Hill Book Co., 1987, p. 1086.

7. Braun, S. R., E. D. Everett, M. C. Perry, and E. V. Sunderrajan, eds. Concise Textbook of Pulmonary Medicine. New York: Elsevier, 1989, pp. 123-133.

8. Warwick, Warren, M.D. Private Communication April 1991.

9. Tharp, G. D. Experiments in Physiology. 5th ed., New York: Macmillan Publishing Co., 1986, pp. 139-144.

10. Taussig, L. M. Cystic Fibrosis. New York: Thieme-Stratton Inc., 1984.

11. Behrman, R. E., V. C. Vaughan, III and W. E. Nelson, eds. The Nelson Textbook of Pediatrics. 13th ed., Philadelphia: W. B. Saunders Co., 1987, p. 926-935.

* Q U I C K C L U S T E R *

Initial Cluster Centers.

Cluster	IMPORTM	CODES	STRAIN	CODEG	FEV	FVC	FEV1
1	7.0000	2.0000	.0000	3.0000	34.1500	115.1700	45.9900
2	1.0000	2.0000	.0000	3.0000	97.1900	28.0000	29.7800
3	1.0000	3.0000	.0000	2.0000	95.6200	148.8700	145.8200

Classification Cluster Centers.

Cluster	IMPORTM	CODES	STRAIN	CODEG	FEV	FVC	FEV1
1	5.2001	2.1776	.1094	2.3684	47.0644	94.7717	51.0831
2	4.0180	2.1010	.1324	2.5074	65.6872	40.8089	27.7645
3	2.3944	2.5666	.0922	1.6892	83.7750	127.9735	118.8138

Final Cluster Centers.

Cluster	IMPORTM	CODES	STRAIN	CODEG	FEV	FVC	FEV1
1	3.5916	2.4754	.1173	1.7556	65.6977	89.1721	65.4064
2	4.0883	2.1943	.1954	2.0497	50.3027	52.0539	29.4637
3	2.7054	2.8085	.0712	1.2900	82.4747	111.3475	100.3203

Number of Cases in each Cluster.

Cluster	unweighted cases	weighted cases
1	1927.0	1927.0
2	906.0	906.0
3	1979.0	1979.0
Missing	7381	
Total	4812.0	4812.0

* Q U I C K C L U S T E R *

Initial Cluster Centers.

Cluster	IMPORTM	CODEG	FEV	FVC	FEV1
1	1.0000	4.0000	84.9100	172.6800	150.4500
2	1.0000	4.0000	64.1000	14.3000	10.0900
3	1.0000	4.0000	35.3500	129.6900	48.5900

Classification Cluster Centers.

Cluster	IMPORTM	CODEG	FEV	FVC	FEV1
1	2.4272	2.1505	80.9432	130.5390	117.6164
2	3.8327	2.5619	41.1350	44.7907	18.8694
3	3.4948	2.5862	56.7131	98.4568	61.2192

Final Cluster Centers.

Cluster	IMPORTM	CODEG	FEV	FVC	FEV1
1	2.3766	1.7042	83.2243	114.4795	104.0333
2	3.3933	2.6035	45.5300	52.8268	26.7922
3	2.9295	2.1171	69.5367	91.0235	70.1553

Number of Cases in each Cluster.

Cluster	unweighted cases	weighted cases
1	1819.0	1819.0
2	1251.0	1251.0
3	2938.0	2938.0
Missing	6185	
Total	6008.0	6008.0

* Q U I C K C L U S T E R *

Initial Cluster Centers.

Cluster	IMPORTM	FEV	FVC	FEV1
1	1.0000	64.1000	14.3000	10.0900
2	1.0000	84.9100	172.6800	150.4500
3	1.0000	35.3500	129.6900	48.5900

Classification Cluster Centers.

Cluster	IMPORTM	FEV	FVC	FEV1
1	3.7929	39.8732	47.3084	19.9954
2	2.5675	80.5924	125.4203	113.1029
3	3.4043	58.6756	95.5309	61.9253

Final Cluster Centers.

Cluster	IMPORTM	FEV	FVC	FEV1
1	3.4043	45.4341	53.0171	26.8405
2	2.4158	82.7902	113.4465	102.6389
3	2.9335	68.9998	90.1445	68.8836

Number of Cases in each Cluster.

Cluster	unweighted cases	weighted cases
1	1259.0	1259.0
2	2013.0	2013.0
3	2736.0	2736.0
Missing	6185	
Total	6008.0	6008.0

* Q U I C K C L U S T E R *

Initial Cluster Centers.

Cluster	CODEG	FEV	FVC	FEV1
1	4.0000	84.9100	172.6800	150.4500
2	4.0000	64.1000	14.3000	10.0900
3	4.0000	35.3500	129.6900	48.5900

Classification Cluster Centers.

Cluster	CODEG	FEV	FVC	FEV1
1	1.9880	80.5924	125.4203	113.1029
2	2.4534	39.8732	47.3084	19.9954
3	2.4640	58.6756	95.5309	61.9253

Final Cluster Centers.

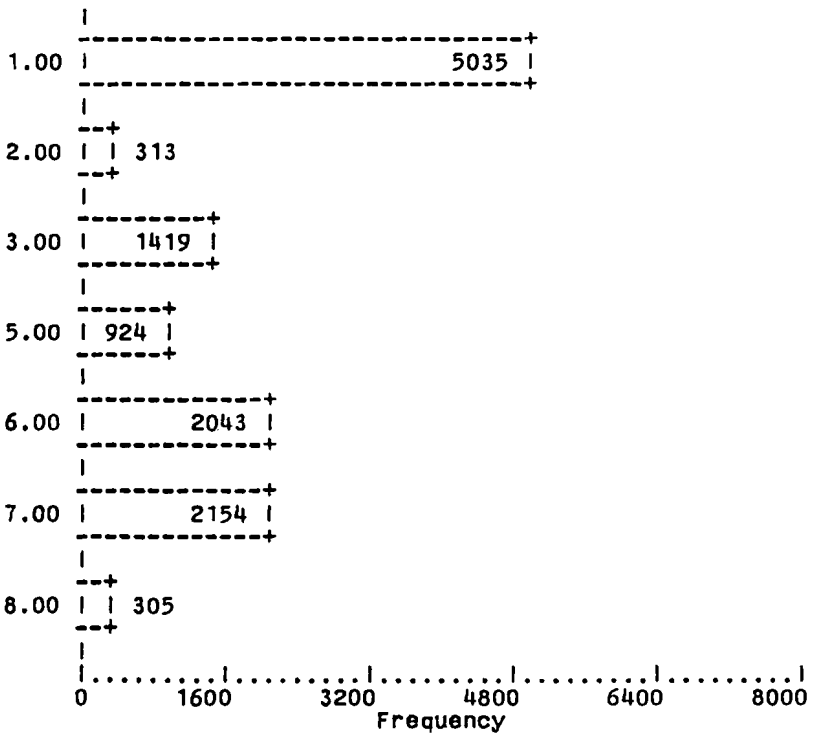
Cluster	CODEG	FEV	FVC	FEV1
1	1.6930	82.7923	113.4438	102.6395
2	2.6029	45.4341	53.0171	26.8405
3	2.1535	68.9982	90.1464	68.8832

Number of Cases in each Cluster.

Cluster	unweighted cases	weighted cases
1	2013.0	2013.0
2	1259.0	1259.0
3	2736.0	2736.0
Missing	6185	
Total	6008.0	6008.0

IMPORTM

Value Label	Value	Frequency	Percent	Valid Percent	Cum Percent
	1.00	5035	41.3	41.3	41.3
	2.00	313	2.6	2.6	43.9
	3.00	1419	11.6	11.6	55.5
	5.00	924	7.6	7.6	63.1
	6.00	2043	16.8	16.8	79.8
	7.00	2154	17.7	17.7	97.5
	8.00	305	2.5	2.5	100.0
	Total	12193	100.0	100.0	



Valid cases 12193 Missing cases 0

- - - - - - - - - - - - - - - P E A R S O N C O R R E L A T I O N C O E F F I C I E N T S - - - - - - - - - - - - - - -

	IMPORTM	CODES	STRAIN	CODEG	FEV	FVC	FEV1
IMPORTM	1.0000 (12193) P= .	.0473 (12193) P= .000	.3779 (10997) P= .000	.2091 (12193) P= .000	-.1880 (6008) P= .000	-.1277 (6008) P= .000	-.1661 (6008) P= .000
CODES	.0473 (12193) P= .000	1.0000 (12193) P= .	-.1564 (10997) P= .000	-.6264 (12193) P= .000	.2864 (6008) P= .000	.2150 (6008) P= .000	.2656 (6008) P= .000
STRAIN	.3779 (10997) P= .000	-.1564 (10997) P= .000	1.0000 (10997) P= .	.2545 (10997) P= .000	-.1830 (4812) P= .000	-.1074 (4812) P= .000	-.1492 (4812) P= .000
CODEG	.2091 (12193) P= .000	-.6264 (12193) P= .000	.2545 (10997) P= .000	1.0000 (12193) P= .	-.2853 (6008) P= .000	-.1984 (6008) P= .000	-.2562 (6008) P= .000
FEV	-.1880 (6008) P= .000	.2864 (6008) P= .000	-.1830 (4812) P= .000	-.2853 (6008) P= .000	1.0000 (6008) P= .	.5927 (6008) P= .000	.8528 (6008) P= .000
FVC	-.1277 (6008) P= .000	.2150 (6008) P= .000	-.1074 (4812) P= .000	-.1984 (6008) P= .000	.5927 (6008) P= .000	1.0000 (6008) P= .	.9071 (6008) P= .000
FEV1	-.1661 (6008) P= .000	.2656 (6008) P= .000	-.1492 (4812) P= .000	-.2562 (6008) P= .000	.8528 (6008) P= .000	.9071 (6008) P= .000	1.0000 (6008) P= .

(COEFFICIENT / (CASES) / 1-TAILED SIG) " . " IS PRINTED IF A COEFFICIENT CANNOT BE COMPUTED