STATISTICAL PROCESS CONTROL IN MEDICAL SURVEILLANCE AN APPLICATION USING SPIROMETRY

THESIS

Paul W. McAree, Captain, USAF

AFIT/GOR/ENS/96M-7

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

AFIT/GOR/ENS/96M-7

STATISTICAL PROCESS CONTROL
IN MEDICAL SURVEILLANCE
AN APPLICATION USING SPIROMETRY

THESIS

Presented to the Faculty of the School of Engineering

of the Air Force Institute of Technology

Air University

In Partial Fulfillment of the

Requirements for the Degree of

Master of Science in Operations Research

Paul W. McAree, M.S., B.S.

Captain, USAF

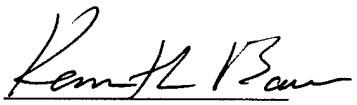
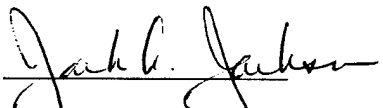
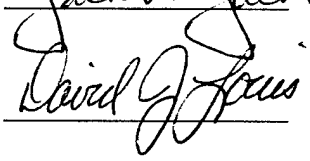
March 1996

Thesis Approval

STUDENT:	Capt Paul W. McAree	CLASS:	GOR96-M

THESIS TITLE: Statistical Process Control in Medical Surveillance
An Application Using Spirometry

DEFENSE DATE: 20 February 1996

COMMITTEE:	NAME/DEPARTMENT	SIGNATURE
Advisor	Lt Col Kenneth Bauer/ENS	
Reader	Lt Col Jack Jackson/ENS	
Reader	Lt Col Dave Louis/SGPO	

Acknowledgements

I would like to thank my committee members: Lieutenant Colonel Ken Bauer, Lieutenant Colonel Jack Jackson, and Lieutenant Colonel Dave Louis. Their help and guidance throughout the research effort were invaluable. Lieutenant Colonel Louis' private library collection simplified the literature review portion of this effort. I would also like to thank the 74th SGPMO for supporting this research effort.

This project was accomplished in conjunction with 2nd Lieutenant Bryan Richardson's research on hazardous work areas with respect to the liver function and I would like to thank him for working together on quite a bit of the project and helping in the writing and presentation portions.

I would also like to thank my brother, Medical Doctor John McAree for reviewing portions of the writing to make sure I was not misapplying statistical techniques to medicine.

Paul W. McAree

Table of Contents

List of Figures / Charts

List of Tables

Abstract

This research effort concentrated on applying statistical process control techniques to the results of seven years worth of spirometry exams of workers at Wright-Patterson AFB in helping the occupational health squadron identify potential work hazard areas. Each spirometry exam was classified as abnormal or normal based on a comparison with established normals or a significant loss of function from the previous year's exam for the individual. Each test was classified into the work area of the individual and the number of abnormalities per population of the work area was examined. This was accomplished through the use of standardized control charts for fraction abnormal. The base population studied was "out-of-control" for all seven years; however, when excluding smokers only one year of the study for the whole base was "out-of-control". Several work areas were identified as being "out-of-control" and recommended for further study by the occupational health squadron.

STATISTICAL PROCESS CONTROL IN MEDICAL SURVEILLANCE AN APPLICATION USING SPIROMETRY

Chapter 1 Introduction

1.0 Background

Wright-Patterson Air Force Base (WPAFB) located in Dayton, Ohio employs approximately 24,000 people of which 15,000 are civilians (26:2). There are about 1,600 different facilities located at WPAFB (26:3) and it is the responsibility of Medical Group's Occupational Medicine Element (74th SGPMO) to ensure each of these occupational areas is not harmful to the people who work in them. A proactive and preventive approach to this goal with respect to occupational pulmonary diseases is desirable.

In light of this goal the 74th SGPMO maintains a health database on personnel working at WPAFB. One to seven years worth of pulmonary function data exists on each of the individuals monitored; the database also includes their work area (building and office), a family history of previous health problems, and personal habits of the individuals (i.e. cigarette smoking and alcohol consumption). However, the 74th SGPMO only maintains this database and does not conduct analysis. The purpose of this research is to conduct the needed analysis to help the 74th SGPMO monitor the different work zones.

The data extracted from the database for this analysis includes the following as identifiers: Social Security Number and Work Zone history. Independent variables

include: Date of Birth, Gender, Race, Height, family history information on Blood Disease, Asthma, Lung Disease, Liver Disease, Hepatitis, and Jaundice, personal information on packs of cigarettes smoked per day, number of years smoked if quit, and total number of years smoked if currently smoking. Also ounces of liquor, bottles of beer, and number of wine glasses consumed per week is included. Possible dependent variables are the results of the pulmonary function test; these include Forced Vital Capacity (FVC), Forced Expiratory Volume in One Second (FEV_1), the Ratio of FEV_1 / FVC, the percent of predicted FVC (FVCpred), the percent of predicted FEV_1 (FEV1pred), and the percent of predicted Ratio FEV_1 / FVC (Ratpred). The number of individuals who have lung data monitored is 1,945 (as of 3 October 1995). The following table shows the breakdown on number of measurements per person in the database.

Table 1-1
Pulmonary Function Tests Measurements

Number of Pulmonary Function Tests	Number of People
1	815
2	469
3	248
4	172
5	104
6	110
7	27

The pulmonary function test is accomplished on a spirometer. A spirometer is "an instrument for measuring air entering and leaving the lungs" (25:1122). The basic principle is that the subject inhales as much air as possible and then forcefully expires the

air as hard as possible as long as possible into the spirometer. The spirometer measures the respective variables during the process. The standard used in this case is three trials and the "best effort" is recorded as the final result. The type of spirometer used in this analysis is a SPIROTECH S400 VER 6C Andersen/Spirotech manufactured by Spirotech, a division of Graseby Andersen.. Figure 1-1 shows a type of spirometer similar to the one used in this study. According to Mr. William M. Yancey, the Pulmonary Function and ECG Technician at the 74th SGPMO, it is approved and does meet the standards set for spirometers set by the American Thoracic Society (ATS). Please see Figure 1-2 for the sample output of such a test. As stated above, the measurements used in this study are FVC, FEV_1, FEV_1 / FVC and the corresponding predicted values. FVC is the maximum amount of air which can be exhaled with a maximum forced effort from a position of maximum inspiration (12:17 & 15:A4). It is the most air a person can breath out after inhaling as much as possible. FEV_1 is the volume of air exhaled in the first second of the FVC performance (12:18 & 15:A7). Both FVC and FEV_1 are measured in liters BTPS, Body conditions: normal body temperature; 37 degrees Celsius, ambient pressure, saturated with water vapor (2:2-3). The ratio is simply the volume of air exhaled in the first second divided by the FVC value and converted to a percentage by multiplying by 100. The predicted values of FVC, FEV_1, and FEV_1 / FVC are calculated by the spirometer based on Knudson normals (which will be defined in Chapter 2).

Normal values and variations among the spirometry results for FVC are based on Gender, Age, Height, Weight, Race, Technical (due to the machine) and Unexplained factors (1:2 & 7:325). Please see Table 1-2. Studies show black adults values of FVC

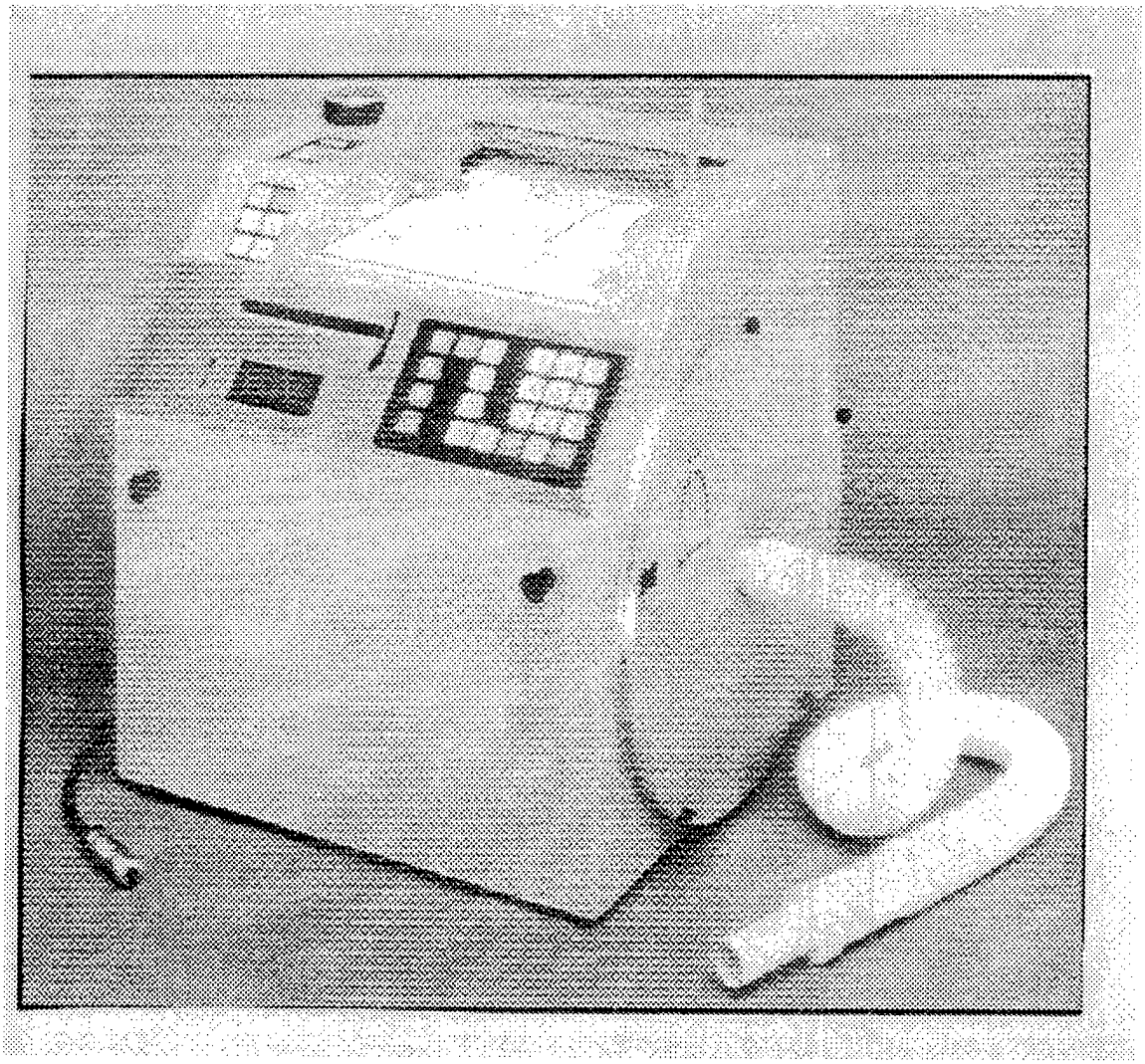


Figure 1-1
A Spirotech Spirometer

and FEV_1 range from 5 to 20% lower than White adults (5:251). Blacks of the same height as Whites, in general, have longer legs and a shorter thorax; therefore these values tend to be lower when controlling for height (10:828). The ratio of FEV_1 / FVC, is also higher in non-whites because the reduction in FEV_1 is less than the reduction in FVC for these subjects (5:250). Technical factors include the instrument being used, the performance of the test administrator, the performance of the subject, interaction between

the administrator and subject, the procedure being used, and the temperature / altitude the test was administered at (5:244). The unexplained factors include past and present health experiences (personal health history, family health history, and current health), past and present exposures (i.e. cigarette smoking and occupational exposures) and socioeconomic factors (7:325).

Table 1-2
The Proportion of Variance in FVC, or FEV_1
(1:2 & 7:325)

Factor	Proportion of Explained Variance
Gender	0.30
Age	0.08
Height	0.20
Weight	0.02
Ethnic Differences	0.10
Technical	0.03
Unexplained	0.27

1.1 Statement of the Problem

Using the information extracted from the health database, accomplish an analysis to determine any signs of possible adverse effects to individual's pulmonary functions which may be a result of their work environment. The scope of this effort is to determine if there is a possible adverse effect between a work zone and the employee's health based on the spirometry results. This analysis is to be used as a screening tool for the 74th SGPMO. If a work zone does not meet "normal" criteria, it is the responsibility of the 74th SGPMO to determine if corrective action is warranted for that zone.

SPIROTECH S400 VER 6C
Andersen/Spirotech
Atlanta, Georgia

ID: 123456789 Time: 9:38 Date: 6/ 8/95
Age: 27 yrs Height: 72 ins
Sex: Male Race: WHITE
22.1'C 760.0 mmHg BTPS: 1.090

Normals: Knudson Cherniack
Best Effort: Test #3 Criteria met

Parameter	Actual	Predicted	%Predicted
FVC	5.66	5.65	100%
FEV.5	3.33	3.56	93%
FEV1	4.40	4.58	96%
FEV3	5.47	5.44	101%
FEF25-75	3.82	5.53	69%
FEF75-85	1.27		
PEFR	10.56	10.25	103%
FEF25	7.35	9.53	77%
FEF50	4.33	6.81	63%
FEF75	1.82	3.58	51%
FEV.5/FVC	.59		
FEV1/FVC	.78	.84	93%
FEV3/FVC	.97		
FET	7.84		

Figure 1-2
Sample Output of Spirometry Test
The Units used are liters BTPS
The Predicted Column is the Knudson Normal

Chapter 2 Literature Review

2.1 Spirometry and Occupational Health

This section will deal with the justification for using the spirometry results for analysis of possible occupational cause of pulmonary disease. The Occupational Health and Safety Administration (OSHA) requires spirometry for people working in certain exposure environments (12:2, 13:229). These include employee exposure to asbestos, coke oven emissions, and cotton dust (14:229). The National Institute for Occupational Safety and Health (NIOSH) recommends pulmonary function testing for exposure to numerous other substances; beryllium, cadmium, formaldehyde, nitrogen-oxides, etc... (14:229). In Albert Miller's "Application of Pulmonary Function Tests" article, he states pulmonary function tests may be the only evidence of certain types of diseases and in the case of obstructive related diseases, they are actually better than a chest x-ray (16:4). Even though the results of the spirometry tests do not establish the diagnosis, they are helpful in suggesting a possible abnormality (16:4). In McKay's article on "Pulmonary Function Testing in Industry", he states spirometry is now regarded as a key part of any respiratory surveillance program (14:229). Not only is spirometry a good instrument with respect to screening for lung disease (4:349), but it can identify pulmonary function abnormalities which may be overlooked otherwise (6:25). Now, why of all the spirometry output shown in Figure 1-1, only three variables are to be used in this study?

In his article "Respiratory Disorders", Dr. Wegman states the most useful information from the spirometry test is the results of the FVC, FEV_1, and the FEV_1 / FVC

ratio when evaluating possible occupational-related respiratory illnesses (24:321). In the same article he also provides a table of major occupational pulmonary diseases and how these three variables react in each of the situations. Please see Table 2-1 for a summary of the different diseases.

Table 2-1 (24:326)
Impact of Disease Type on Variables

Type	Occupational Example	FEV1	FVC	Ratio FEV1/FVC
Restrictive	Silicosis, Asbestosis	Decrease	Decrease	No Change
Obstructive	Byssinosis, isocyanate asthma	Decrease	No Change	Decrease
Granulomata	Beryllium disease	Decrease	Decrease	N/A
Pulmonary edema	Cadmium Poisoning	Decrease	Decrease	N/A

Restrictive lung disease means the lungs are too stiff and/or too small. An individual with restrictive lung disease will slowly experience difficulty in breathing (9:175). An obstructive lung disease (a.k.a. Chronic Obstructive Pulmonary Disease) means the airways are becoming blocked and/or the lungs are losing their structural integrity. With respect to occupational asthma, the airways become inflamed due to the irritant and this results in blocking (9:145). Granulomata indicates the presence of lung nodules. A pulmonary edema is basically fluid in the lungs.

The ATS reinforces Dr. Wegman's view by stating the FVC, FEV_1, and FEV_1 / FVC ratios are the three basic factors in the interpretation of a spirometry exam and goes on to state in diagnosing for an obstructive disease, the FEV_1 / FVC ratio is the most important measurement (1:1212). Several studies on occupational lung disease utilized these three variables; therefore setting precedence for the approach taken in this research. Parker's study on the pulmonary function of autobody repair workers was based on measurements of FEV_1, FVC and the FEV_1 / FVC ratio; the FEV_1 / FVC ratio results were used to show indications of possible obstructive disease (22:768-771). Monson cites a study of Potash mine workers in Canada in which the pulmonary function was measured by spirometry. He showed how the FEV_1 value related to SO_2 exposure (19:215-216).

2.2 Defining "Abnormality" With Respect to a Reference Population

As stated in Chapter 1, general population normal values for FVC, FEV_1, and the FEV_1 / FVC ratio exist with respect to the factors sex, age, height, weight, and race. The analysis of a specific population is possible because of these general population normal values (19:214). The spirometer used in this study outputs a percent of predicted value for each person based on the Knudson method. Table 2-2 shows the coefficients for the regression equations the Spirotech spirometer uses in predicting the normal values based on the equation:

$$y = C + [\text{age(years)} * \text{age coefficient}] + [\text{standing ht(cm)} * \text{ht coefficient}] \quad (13:589)$$

where y is either the predicted FVC value, the predicted FEV_1 value, or the predicted FEV_1 / FVC ratio value. For blacks, the predicted values for FVC and FEV_1 need to be multiplied by 0.85 to account for the racial difference (10:848).

In the article "Predictable Confusion", Dr. Glindmeyer states one set of regression equations could show a person to be normal with respect to their lung function, while another set would show the person to be abnormal (10:845). He recommends using the values of Knudson because they are based on a large sample of healthy, never-smokers, who are free from any respiratory symptoms (10:848). This is consistent with the recommendations of OSHA which adopted the predicted normals based on Knudson's testing for its 1978 cotton dust standard (12:2).

Table 2-2
Coefficients for Knudson Regression Equations (13:590)

Test	Gender	Age	C Coefficient	Age Coefficient	Ht Coefficient	Std Dev of Prediction
FVC	Male	< 25 Years	-5.508	0.078	0.050	0.544
FVC	Male	≥ 25 Years	-5.459	-0.029	0.065	0.601
FVC	Female	< 20 Years	-3.469	0.092	0.033	0.500
FVC	Female	≥ 20 Years	-1.774	-0.022	0.037	0.519
FEV_1	Male	< 25 Years	-4.808	0.045	0.046	0.523
FEV_1	Male	≥ 25 Years	-4.203	-0.027	0.052	0.541
FEV_1	Female	< 25 Years	-2.703	0.085	0.027	0.422
FEV_1	Female	≥ 25 Years	-0.794	-0.021	0.027	0.434
FEV_1/FVC	Male	All Ages	103.64	-0.140	-0.087	6.721
FEV_1/FVC	Female	All Ages	107.38	-0.109	-0.111	7.664

In Bascom and Ford's article, "Don't Just 'Do Spirometry'- Closing the Loop in Workplace Spirometry Programs", the answer to the question, "Is the output normal?", is answered by the comparison of the individual test result with the predicted value based on the regression equation used (4:355). The general cutoff for abnormals recommended for FEV_1 and FVC is less than 80% of the predicted value (4:355). The standard deviations about the regression line for these predictions are shown in Table 2-2. In Miller's "Prediction Equations and 'Normal'" article, he cites the American Lung Association's handbook on *Chronic Obstructive Pulmonary Disease* recommendation that 80% of predicted is abnormal with respect to both the FEV_1 and FVC variables (17:203). With the prediction equations, several organizations set guidelines for determining abnormality of the pulmonary function with respect to the predicted values. The NIOSH, ATS, and American Medical Association (AMA) recommend the lower limits of Normal as shown in Table 2-3.

Table 2-3 (15:C-7)
Recommended Lower Limits of Normal

	FVC (pct of predicted value)	FEV1 (pct of predicted value)	FEV1 / FVC Observed Value
NIOSH	80	80	70
AMA	80	80	70
ATS	80	80	75

The American Thoracic Society does recommend another approach to defining an abnormal reading with respect to the population predicted values. They recommend normal ranges based on fifth percentiles of the reference population, which is considered to be below the lower limit of normal (1:1206) Miller agrees with this practice because having a fixed percent predicted value will cause false negatives; classification of "normal" readings as "abnormal" (17205). Dockery adds that the use of 80% of the predicted value as a lower limit will more likely misclassify shorter and older people as "abnormal" and also misclassify taller and younger people as "normal" (7:326). This is because the variation by height and age is not consistent. McKay acknowledges the 80% lower limit for normal value, but believes a 95th percentile method is more valid (14:232) because this method will have fewer false positives. Table 2-4 summarizes the abnormality cutoff (the lower fifth percentile) with respect to the general population normal values using the Knudson regression equations and the percent of predicted values.

Table 2-4 (15:C-11,12)
Lower Fifth Percentiles Based on Knudson
Values are Percent of Predicted

Gender	Age	FVC	FEV_1	FEV_1 / FVC Ratio
Male				
	25-39 years	81.1%	79.1%	87.0%
	40-84 years	73.4%	77.2%	87.0%
Female				
	20-39 years	76.9%	70.3%	85.4%
	40-70 years	75.2%	77.9%	85.4%

2.3 Defining Abnormal Declines of Lung Function

The database does have multiple readings of individuals, therefore a longitudinal study can also be accomplished with respect to this data. According to Dr. McKay, even though we can compare a person's test results to the reference values, comparing them to previous tests is more desirable (14:234). This is because the coefficient of variation of a given test within the individual is smaller than the population's coefficient of variation (14:234). The coefficient of variation is the standard deviation as a proportion of the mean of the population. Where the population's variation is based on all the factors listed in Table 1-2, the individual's variation will be limited to the technical and unexplained factors of taking the spirometry exam itself. Bascom and Ford state using the person's own longitudinal data is much better than comparing a single value versus a population predicted value (4:357). In the event of having multiple test results letting the person be their own control is much more desirable (12:29) Table 2-4 shows the percent changes before a meaningful diagnosis of abnormality for a longitudinal study can be claimed.

Table 2-4 (15:H-6,7)
Abnormality in a Longitudinal Study

Source	FVC	FEV_1	FEV_1 / FVC Ratio
ATS	Annual Decline of 15% or more	Annual Decline of 15% or more	
NIOSH	Annual Decline of 10% or more	Annual Decline of 10% or more	Annual Decline of 5% or more

2.4 Summary of Classification

From the information stated above, the cuts for an "abnormality" classification with respect to the result of the spirometry test will be a combination of the reference population normals using 5% of the lower limit (based on the values in Table 2-3) for FVC and FEV_1 and a longitudinal study where multiple tests on individuals is available (based on the values in Table 2-4). For ratio the observed value of 0.70 or lower will be used and a longitudinal decline of 10% or more. The 5% cutoff yielded too many false abnormalities. If the person falls into *any* of these *categories* their *classification* will be *abnormal*. Dr. Glindmeyer states a subject should be classified as abnormal if he/she is below the normal range or exhibits an abnormal longitudinal decline, even if it is within the normal range (12:849). Another reason for combining the two methods in this study is the information is available and if the person enters the work environment with a well above predicted value and declines at 20% per year, a study with respect to only the reference population values would not classify this person as abnormal until it is too late for the screening purpose of the study to be useful (4:357).

2.5 Cigarette Smoking

No epidemiological study on pulmonary disease can be accomplished without accounting for cigarette smoking (1:1206, 19:214). It is definitely a potential confounding factor in the analysis of possible pulmonary disease for a work group area (1:1206). Cigarette smoking does lead to airflow obstruction, which will show in a lower FEV_1 value (6:25). Even if a smoker has one "normal" pulmonary function test, he/she should

not believe they are free of problems; their loss of lung function will show in the future (6:25). Almost all of the literature states smoking does need to be accounted for in this study. The ATS states it does need to be an independent variable in the analysis, but the most appropriate method of measurement is not known ; i.e. binary yes/no variable, amount currently smoked, pack-years, etc.(1:1206). The three variables in the database are: current smoking status, past smoking status, and number of packs smoked per day.

2.6 Statistical Process Control

Each observation of an individual will be classified as "normal" or "abnormal" with respect to the FVC, FEV_1, and the ratio FEV_1 / FVC. A table showing the logic for the classifications is presented in chapter 3. These classifications are the attributes of the analysis. The statistical quality technique which corresponds to the fraction of abnormal occurrences per work zone is a control chart for fraction nonconforming; a p chart (20:147). According to Montgomery, the item may have several quality characteristics (in this case each of the individual FVC, FEV_1, and ratio FEV_1 / FVC readings as well as the longitudinal evaluations) which may be classified as abnormal. If any of these are abnormal, the subject is classified as abnormal (20:148). The fraction defined as abnormal would then be the ratio of the number of people in each work area classified as abnormal divided by the total number of people in that work area (20:148). Based on this fraction, a standard value of p decided by management (the 74th SGPMO) can be used to develop the upper control limit (UCL), the center line, and the lower control limit (LCL) in a control chart (20:149). The UCL and the LCL define the range in which the fraction

classified as abnormal would be accepted, or in other words, the work area would be considered in control with respect to lung abnormalities. The equations for the UCL, center line, and LCL follow:

$$UCL = p + 3\sqrt{\frac{p(1-p)}{n}} \quad (20:149)$$

$$Center\ line = p \quad (20:149)$$

$$LCL = p - 3\sqrt{\frac{p(1-p)}{n}} \quad (20:149)$$

In the above equations 'p' is the fraction nonconforming decided by management and 'n' is the number of individuals tested in the work area. When the fraction of abnormalities is plotted for each work area, the points which fall above the UCL identify the work areas for further investigation to determine if there is an assignable cause (20:153). In the case where sample sizes vary (which they do for the number of people per work area) the simple approach is to base the control limits on each respective sample size (20:163). This is the method employed in this study. For presentation purposes the "standardized" control chart is used. This plots the same information as the p chart; however, in standard deviation units (20:167). The UCL is 3 standard deviation units and the variable which is plotted is Z and is found by the following equation:

$$\hat{Z} = \frac{(\hat{p} - p)}{\sqrt{\frac{p(1-p)}{n}}} \quad (20:149)$$

where '$\hat{Z}$' is the Sample Standard Deviation, '$\hat{p}$' is the sample percent abnormal, and 'p' is the Center Line from above.

Chapter 3 Methodology

3.1 Overview

The steps in transferring the database from the Phoenix system to the SAS system, the logic used in creating a workable database, and the programs used in the analysis will be outlined in this chapter. This thesis effort was accomplished in conjunction with the same type of research with respect to liver tests, therefore the logic for creating the database including the liver data is also presented. Figure 3-1 depicts the process flow in the creation of a SAS database, the classification of abnormalities, and the implementation of control charts in which the lung and work zone data can be analyzed.

3.2 Creating the SAS Database

The first step involves downloading the appropriate data elements from the Phoenix Health database system to flat ASCII files. This was accomplished using the Sentry Health Surveillance system (a menu driven data inquiry system for the Phoenix database). Due to the structure of this database the following 7 output files were created:

Thesis1:(Information on Date of Birth, Sex, Race, Height, Date of Exam)

ssan	dob	sex	race	height	examdate
007645826	1961	M	C	72	1995/05/26

Thesis2:(Lung Data Information)

ssan	date	fev1	fvc	ratio	fvcpct	fevpct	ratiop
002547331	1993/11/17	3.01	3.61	0.83	98	100	98

Thesis3:(Liver Data Information)

ssan	sgpt	sgot	date	ap	ggt	bilirubin	albumin
005527773	25	27	1989/11/06	96	106	0.5	4.3

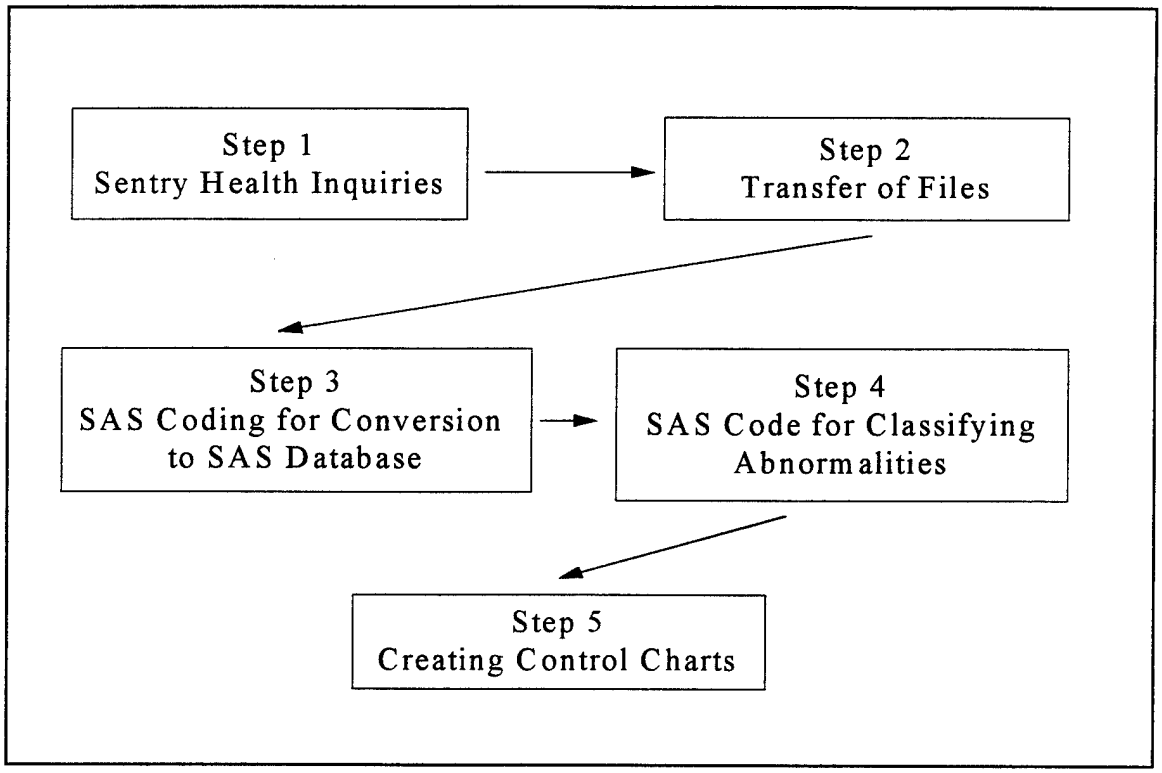


Figure 3-1
Methodology

Thesis4:(White Blood Count Data Information)

ssan	wbc	date wbc	hematocrit
002547331	7.2	1993/05/13	38.9

Thesis5:(Work Zone Information)

ssan	zone	start date	end date
002485546	B640D3	1993/02/08	1994/04/12

Thesis6:(Family History Information on Blood Disease, Asthma, Lung Disease, Liver Disease, Hepatitis, and Jaundice; Personal Information on Packs of day smoked, Number of years, and how long since they quit)

ssan	bld	ast	lun	liv	hep	jau	smo	smo	smo
005208011	B	B	B	B	B	B	2	20	30

A "B" entry means no history and an "A" entry means there is a history.

Thesis7:(Personal information on ounces of liquor, bottles of beer, and number of wine glasses consumed per week)

ssan	d1	d2	d3
033420318	2	2	3

As each exam is repeated, a new entry based on the SSAN is created; therefore we do have multiple entries on the same person. After the downloading of each respective file to floppy disks, the files are transferred to the UNIX mainframe system at AFIT via the WS-FTP protocol.

The third step in this process is creating a SAS database of the above information in a workable format. There are several steps involved in this process. The first one is making a SAS database out of the above files. This is accomplished in the CONVERT.SAS program (please see appendix I). After each of the databases are created, they are then converted back to flat ASCII files so that the information (FVC, FEV_1, SGPT, etc.) is on the same line as the SSAN. In this step, the multiple observations per individual is done away with and there now exists flat files where all the information per SSAN is on one line. This is accomplished in the *RAW.SAS programs (please see appendix II). After each *RAW.SAS program runs, the output file is saved as a flat ASCII file. The next step involves reading back in the raw files into SAS databases and making one dataset for the entire information. This is accomplished in the MERGEALL.SAS program (please see appendix III). This creates a database with all the lung, liver, and zone information which is called "HEALTH.WPAFB2".

3.3 Classification of Common Work Zone Exposure Areas

The work zones are based on the area of base in which a person works; this can be either A, B, C, or K, the building number in which they work, a letter for identifying

common exposures, and a number for furthur breakdown of the exact area within the building they work. For purposes of this analysis, common exposure areas for work zones will be based on the area, building number, and first letter of exposure.

3.4 Classification of Abnormal Lung Function Test

Step four involves classifying each of the individual FVC, FEV_1, and FEV_1 / FVC readings as either normal or abnormal. The LUNGYRS1.SAS and LUNGYRS2.SAS programs (please see appendix IV) accomplish this task while also ensuring the correct classification of zone for each observation is made. The logic for classifying a test as abnormal based on the literature review in chapter 2 is summarized in Table 3-1.

Table 3-1
Logic Table for Classification of Abnormalities

Test	Gender	Age	Value
Pct Predicted of FVC	Male	< 40 years	81.1 or less
Pct Predicted of FVC	Male	> 39 years	73.4 or less
Pct Predicted of FVC	Female	< 40 years	76.9 or less
Pct Predicted of FVC	Female	> 39 years	75.2 or less
Pct Predicted of FEV_1	Male	< 40 years	79.1 or less
Pct Predicted of FEV_1	Male	> 39 years	77.2 or less
Pct Predicted of FEV_1	Female	< 40 years	70.3 or less
Pct Predicted of FEV_1	Female	> 39 years	77.9 or less
Observed Ratio	Both	All Ages	70 or less
Annual Decline of FVC	Both	All Ages	15% or more
Annual Decline of FEV_1	Both	All Ages	15% or more
Annual Decline of Ratio	Both	All Ages	10% or more

The following assumptions and deletions were made:

- Only one pulmonary function measurement per year per SSAN is allowed
- If the date of the first pulmonary function test is missing and a second test exists, it is made exactly one year before the date of the second test.
- If date of birth was a missing value, the subject was classified as over 40 years
- If gender was missing, the subject was classified as a Male
- If the subject had multiple lung exams in one year, only the last exam is counted

3.5 Creating Control Charts

The output data from the LUNGYRS2.SAS program will be saved to a raw file and imported into a Microsoft Excel Spreadsheet. The Spreadsheet database has the format displayed in Table 3-2.

After the data is input into spreadsheet format, *p*-charts and standardized control charts will be developed using the formulas in chapter 2. Two main charts will be created. The first one will be a *p*-chart with a center line of 0.10 measuring the percent of abnormalities for all people in all zones. The second chart will also be a *p*-chart with a center line of 0.10. This chart will exclude all people who smoke and will only measure abnormality rates of nonsmokers in each respective zone. A column included in the spreadsheet will identify the case numbers which fall above the Upper Control Limit indicating an "out-of-control" process. This information will then be summarized and given to the 74th SGPMO for action.

Table 3-2
Excel Spreadsheet Database

Data Name	**Description**
CASE	Zone Identifier by Case Numbers
ABNML2	Total Number of Nonsmoker Abnormalities in Zone
ABNORMAL	Total Number of Abnormalities in Zone
COUNT	Total Number of People in Zone
COUNT2	Total Number of Nonsmokers in Zone
SAB	Total Number of Smoker Abnormalities in Zone
SMOKE	Total Number of Smokers in Zone
YEAR	Year of Test
ZONE	Work Area Identifier
CONTROL	1 if out of control, 0 if in control

Chapter 4 Results

4.1 Overview of Chapter

The results of this analysis effort will be summarized in this chapter. The first part will include a summary of the number and type of abnormalities for each of the years 1989 through 1995. The second part will include standardized control charts and *p*-charts to graphically indicate which work areas and years are considered above an acceptable limit with respect to the number of abnormal lung results. This section will also include a summary of the work areas which were consistently out of control.

4.2 Summary of Abnormal Classifications

There exists up to six possible ways for an individual's lung function test to be classified as abnormal using the logic in Table 3-1. The FVC or FEV_1 may be less than a certain percentage of a predicted value, the FEV_1 / FVC ratio may be less than a certain percentage, or any of the three measurements may have a larger than expected normal decline from the previous year's measurement. Table 4-1 shows the number of people with at least one abnormal reading from the six possibilities, the total number of individual tests' administered in each year, the number of nonsmokers, the number of nonsmokers with at least one abnormal reading and the percentage of abnormal readings by smokers and nonsmokers for each year of the study. An interesting fact is that the percentage of nonsmokers with an abnormal lung function test is lower than the population group as a whole for each of the years under study.

Table 4-1
Summary of Abnormal Tests by Year

Year	Total	Number Abnormal	Pct Abnormal	Total Nonsmokers	Number Abnormal Nonsmokers	Pct Abnormal Nonsmokers
1989	232	59	25%	94	18	19%
1990	923	215	23%	443	96	22%
1991	854	137	16%	420	50	12%
1992	733	105	14%	311	31	10%
1993	836	127	15%	379	45	12%
1994	676	104	15%	316	31	10%
1995	290	54	19%	127	16	13%

Table 4-2 shows the frequency of the abnormal "hits" for an individual lung test for each year of the study. The range of values is from one to six as described above. Note: 1989 has a range from one to three because it is treated as year one in the study and can have no abnormal "hits" for a longitudinal decline.

Table 4-2
Summary of Abnormal Readings for One Lung Function Test

Year	One	Two	Three	Four	Five	Six
1989	33	22	4	N/A	N/A	N/A
1990	96	86	25	7	1	0
1991	68	46	15	6	2	0
1992	50	39	7	4	4	1
1993	65	42	16	3	0	1
1994	65	29	7	2	1	0
1995	27	21	5	1	0	0

Related to Table 4-2 is the information presented in Table 4-3 which shows for each lung function where the abnormal "hits" took place for each of the years.

Table 4-3
Classification of Abnormal Readings

Year	Abnormal FVC	Abnormal FEV_1	Abnormal Ratio	Abnormal Decline of FVC	Abnormal Decline of FEV_1	Abnormal Decline of Ratio
1989	48	32	9	0	0	0
1990	146	116	38	27	24	25
1991	52	77	46	10	13	41
1992	50	60	38	18	11	14
1993	54	63	48	18	8	24
1994	27	49	43	6	10	22
1995	25	28	19	11	3	2

From the information gathered by classifying each individual function test as normal or abnormal based on the six criteria, the next step is creating control charts to analyze which work areas may be out-of-control.

4.3 Control Charts

Chart 4-1 is a standardized control chart based on the number of people with an abnormal lung result versus the total number of people in each of the work areas for the respective years. A work area was analyzed only if more than **5** individual pulmonary function tests were administered for the particular year. Due to the large number of area-year categories studied case numbers are assigned to identify each area-year on the graph. The interpretation of each case number can be found in appendix V. A p value of 10%

was determined to be used in the application of the control chart. This is used because if just FVC and FEV_1 were measured and assumed independent, 9.75% of the normal tests would be falsely classified as abnormal.

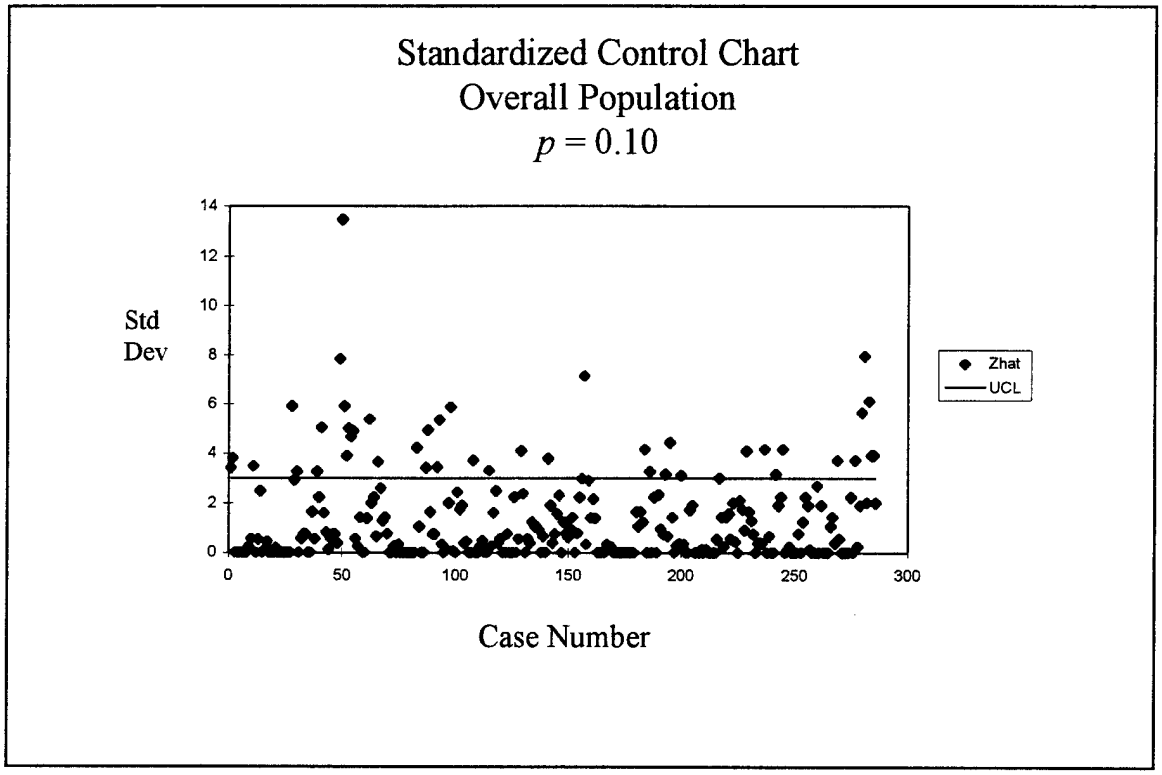


Chart 4-1
Overall Population Standardized Control Chart

All of the points above the 3 standard deviation Upper Control Limit are considered out-of-control work-year areas. There are 44 such areas of a possible 286 work-year combinations. Seven of these areas include the overall base average for each of the years 1989-1995. Table 4-1 above showed the overall abnormal percentage for each of these areas; all well above 10%. The actual standard deviations above a 10% center line ranged from 3.9 to 13.5 for the overall base as displayed on the standardized control chart above. A summary of the 44 zone-years out-of-control is shown in Table 4-4.

Table 4-4
Out-of-Control Work Areas

Zone	Years Out of Control
A1400	89, 90
A278A	90
A830F	89, 93
A867A	90
A876C	90
Whole Base	89, 90, 91, 92, 93, 94, 95
B36A	90, 94
B470B	91
B490A	90, 91
B4B1	91, 92
B4D	90
B5E	92
B5J1	95
B65A	90
B745B	91
B770A	90
C13D	93
C13F	90
C13R	94
C163A	90, 95
C22A	90
C22I	93
C29A1	89, 94
C4020A	90
C883A	94
C91B1	94
K1240	89, 90, 92, 93, 94

Due to the fact the percentage of abnormal readings for nonsmokers was much lower than the overall average for the base, a standardized control chart for nonsmokers only also was developed. This chart is shown in Chart 4-2. The number of out-of-control work area-years in this case dropped from 44 to 15. The overall base average for nonsmokers has only one year out-of-control (1990).

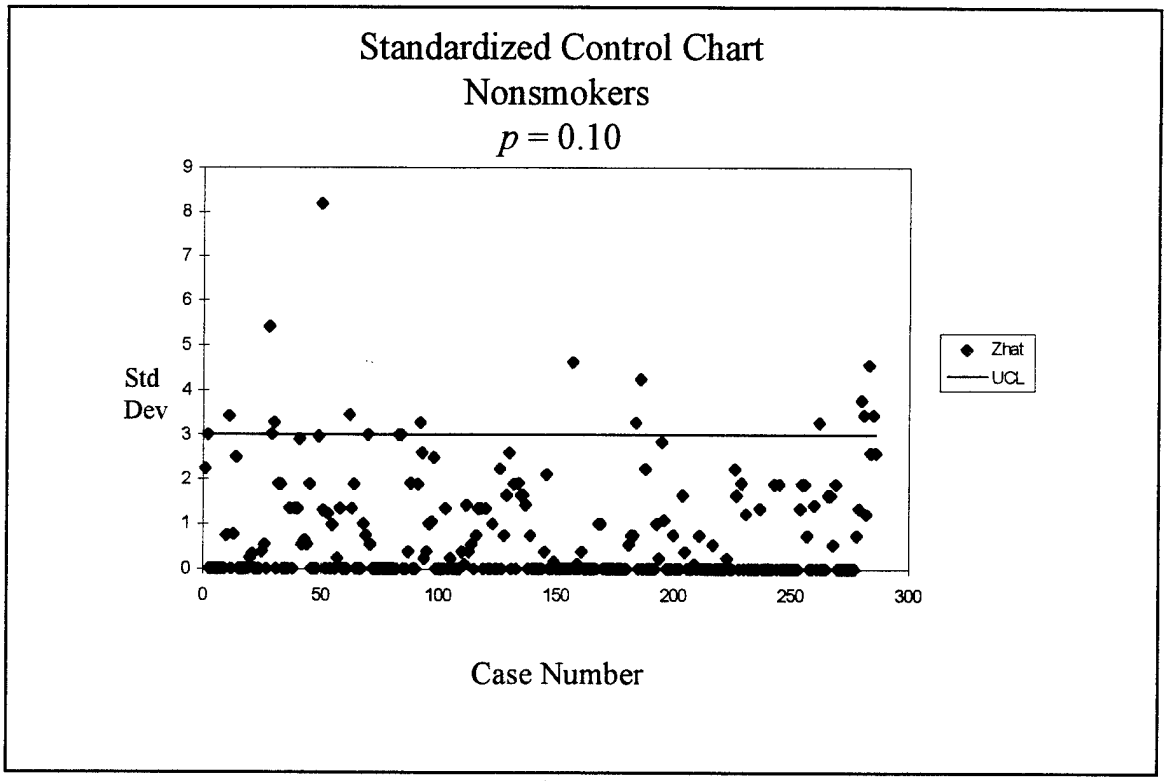


Chart 4-2
Nonsmokers Standardized Control Chart

Table 4-5 shows the work areas and given years for which they were out of control when smokers were excluded from the study.

Table 4-5
Nonsmoker Out-of-Control Work Areas

Zone	**Years Out of Control**
A278A	90
A830F	89, 90, 93
Whole Base	90
B36A	90
B4B1	91
B770A	90
C13D	93
C13F	90
C4067A	93
K1240	89, 90, 92, 94

Zone K1240 is prevalent in both cases. This work zone is the Heat Plant located in the Kittyhawk area of the base. A *p*-chart for this specific zone is shown in Chart 4-3. A *p*-chart of zone K1240 also is shown for the non-smoking workers only in Chart 4-4.

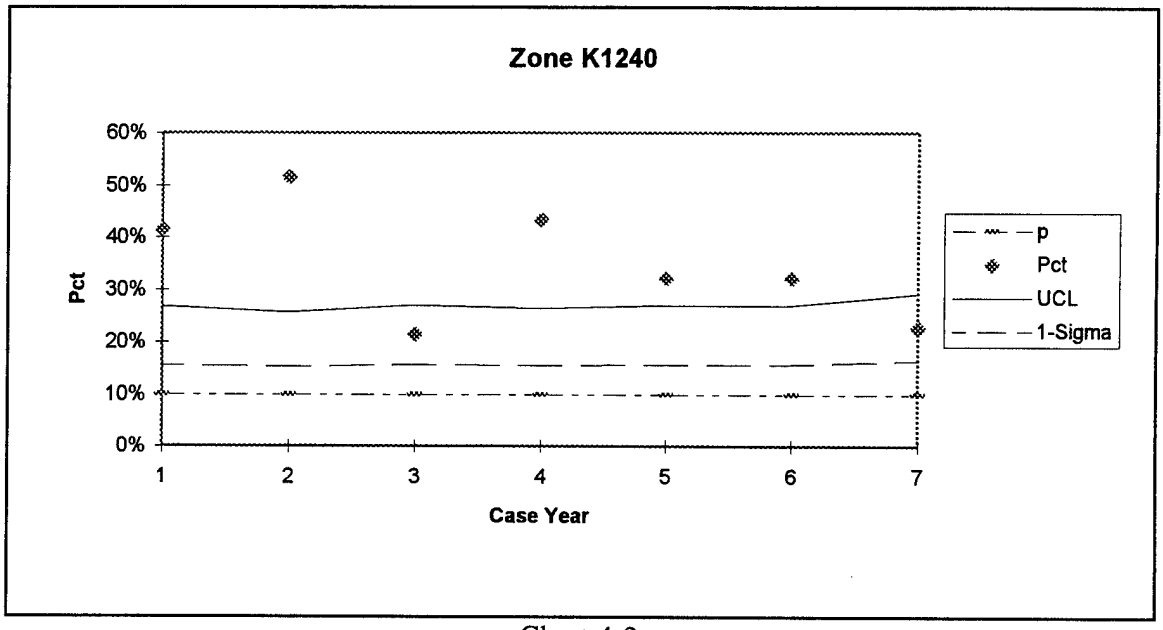


Chart 4-3
Zone K1240

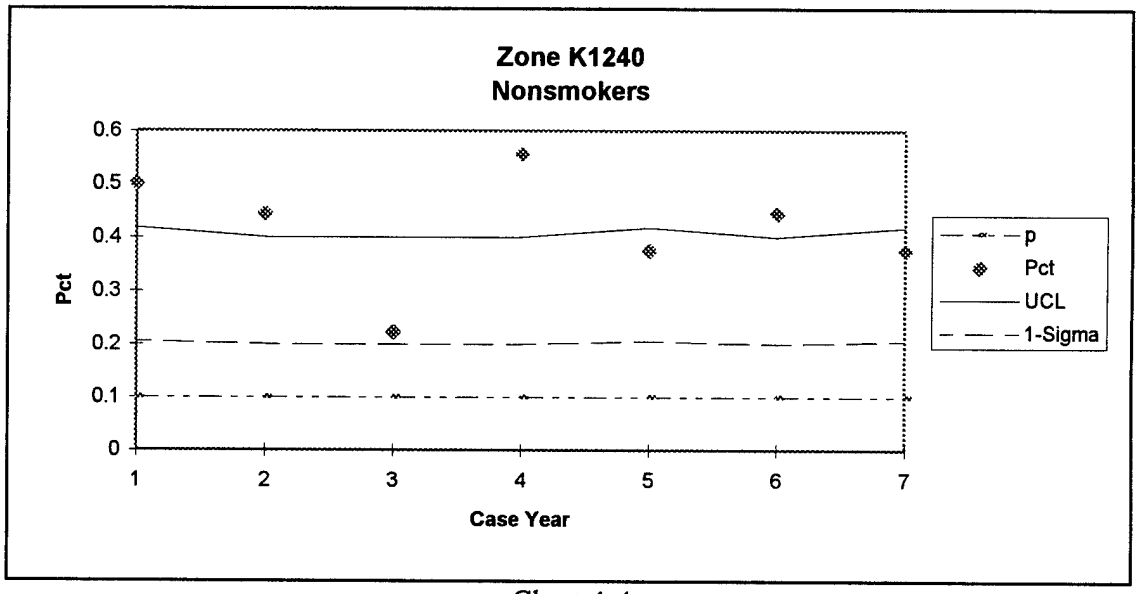


Chart 4-4
Zone K1240 Nonsmokers

As seen in the above charts the lowest percentage of abnormalities is over 20% (in all seven years the percent of abnormalities are at least one-sigma above the *p*-value). In both cases: including and excluding the smoking population this work area has an abnormal rate which is well above the one-sigma limit in all years of the study. This is a zone which is recommended for further study by the 74th SGPMO.

Montgomery states a process is considered out-of-control if 2 of 3 consecutive points are above the 2-sigma limits or 4 of 5 are above the 1-sigma limit (20:117). Based on this criteria other zones recommended for further study based on the overall population are: B36A (Heat Distribution Center in Area B) which had 6 of the 7 years at least one standard deviation above the *p*-value, B770A (Heat Plant in Area B) which also had 6 of the 7 years at least one standard deviation above the *p*-value, C22I (88th Civil Engineering Squadron's Project Painters) which had 4 of the 5 years at least one standard deviation above the *p*-value, and zone C163A (Fire Stations #1, #2 and #5) which had 2 of the 7 years at least 3 standard deviations above the *p*-value. Charts 4-5 through 4-12 show the results of these zones including and excluding the smoking population. Appendix VI reflects the Excel Spreadsheets used in creating the charts for each of the respective areas.

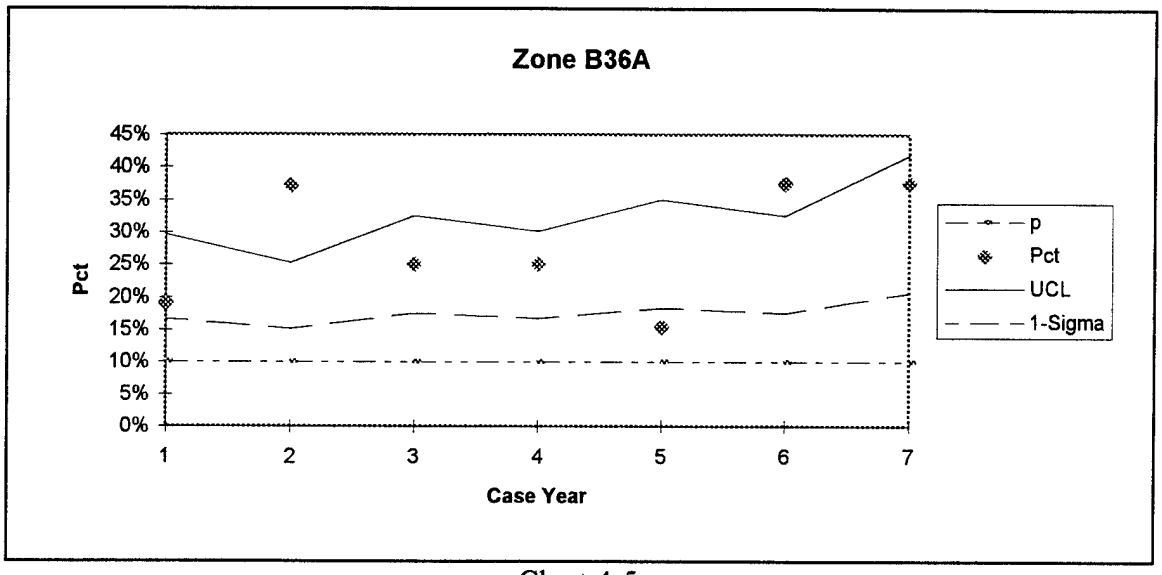


Chart 4-5
Zone B36A

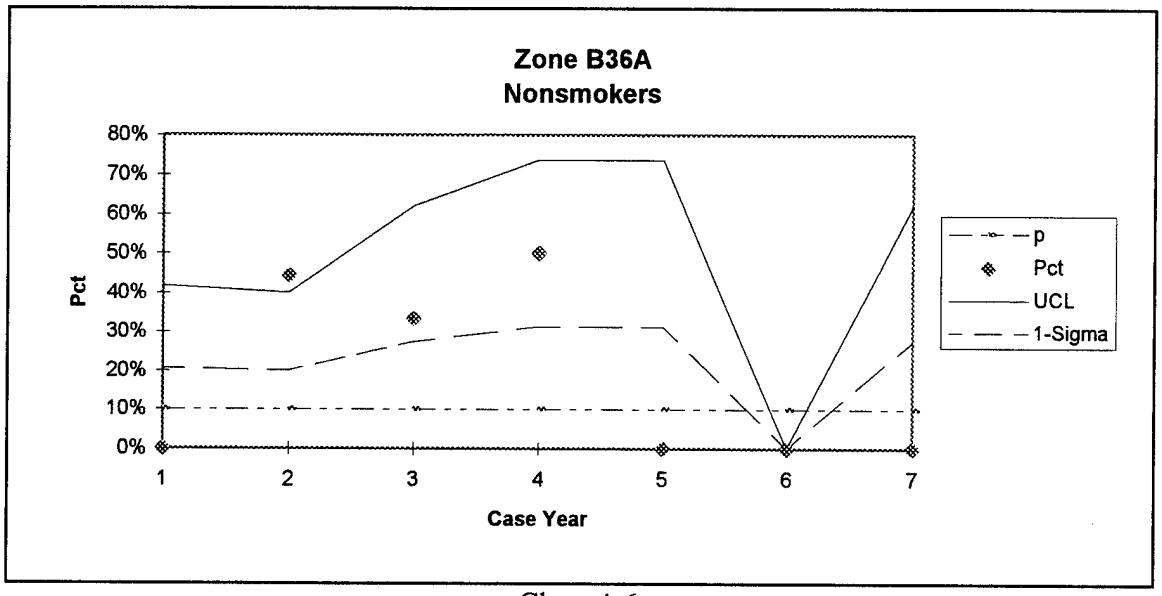


Chart 4-6
Zone B36A Nonsmokers

Chart 4-5 shows zone B36A as having consistently high abnormal rates. By excluding people who smoke and accomplishing the *p*-chart of Chart 4-6, we see this zone actually has no abnormalities for the past 2 years in which nonsmokers were measured (case year 6 had zero nonsmokers measured).

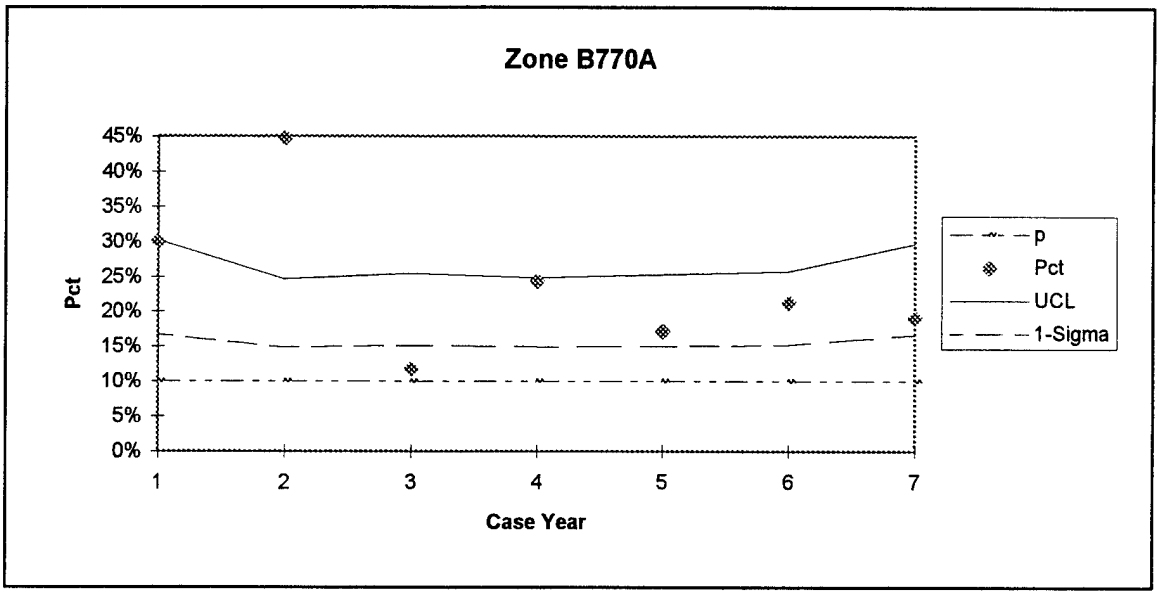


Chart 4-7
Zone B770A

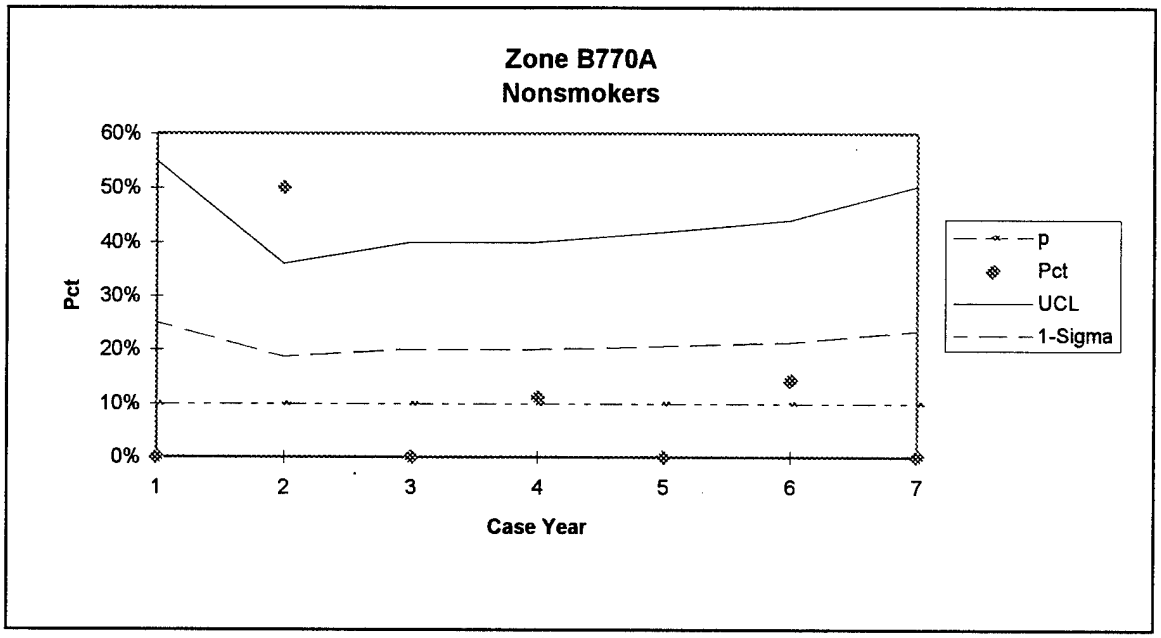


Chart 4-8
Zone B770A Nonsmokers

Chart 4-7 shows zone B770A as having consistently high abnormal rates. By excluding people who smoke we see this zone actually has no abnormalities above the one-sigma limit since case year 2 of the study.

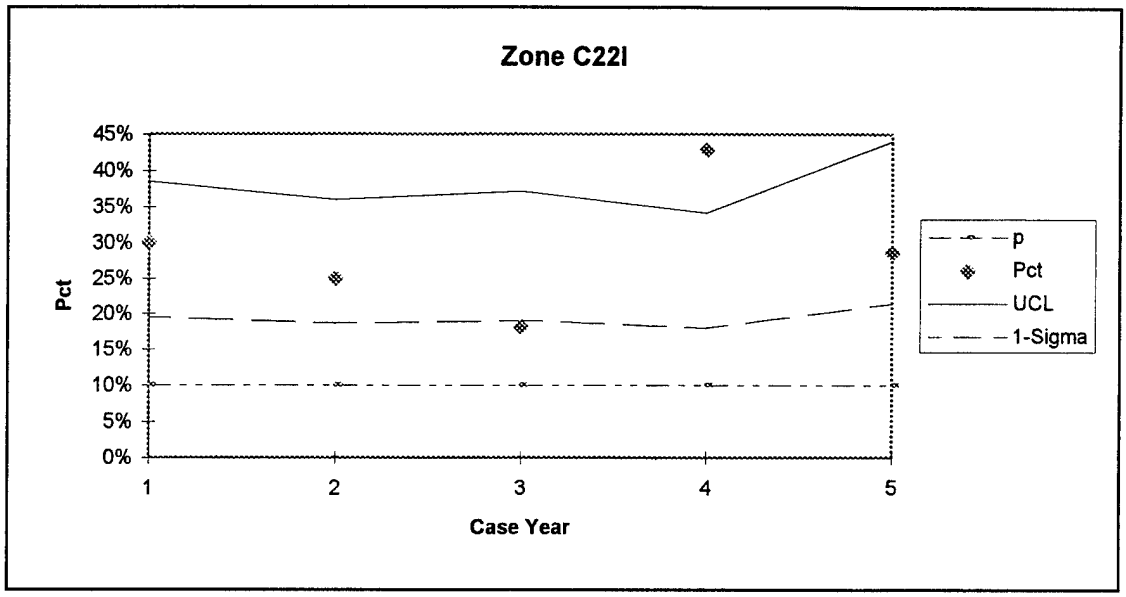


Chart 4-9
Zone C22I

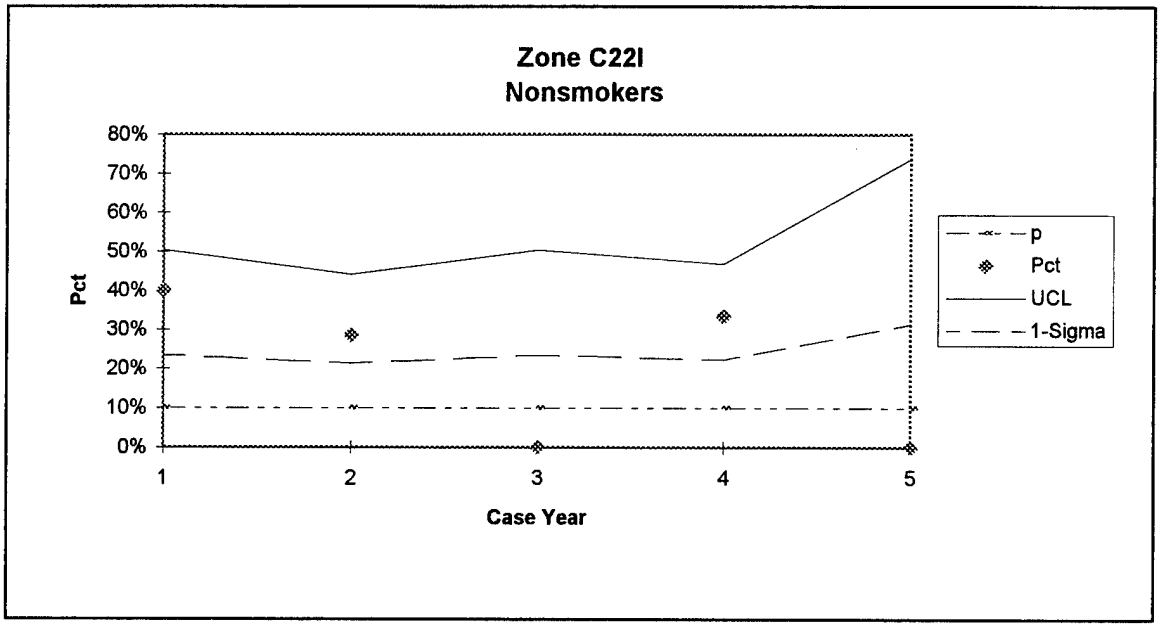


Chart 4-10
Zone C22I Nonsmokers

Chart 4-9 shows zone C22I as having consistently high abnormal rates. After excluding people who smoke, this zone still has a higher than 1-sigma percentage of abnormalities for three of the five years. Case year 1 is 1990 and Case year 5 is 1995.

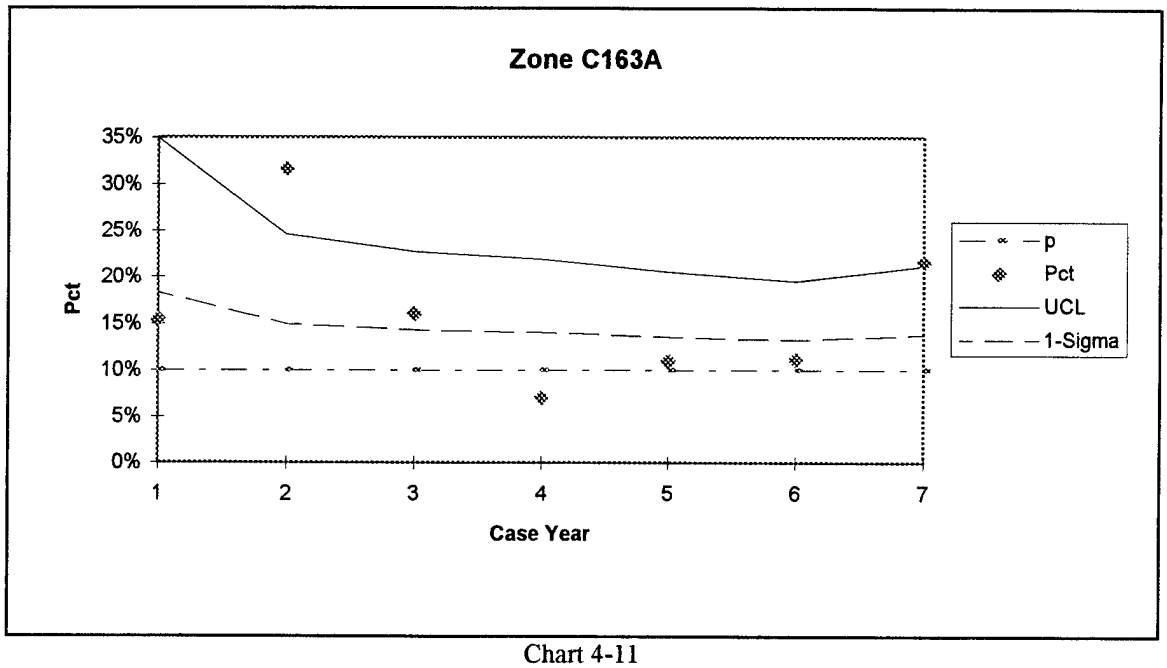


Chart 4-11
Zone C163A

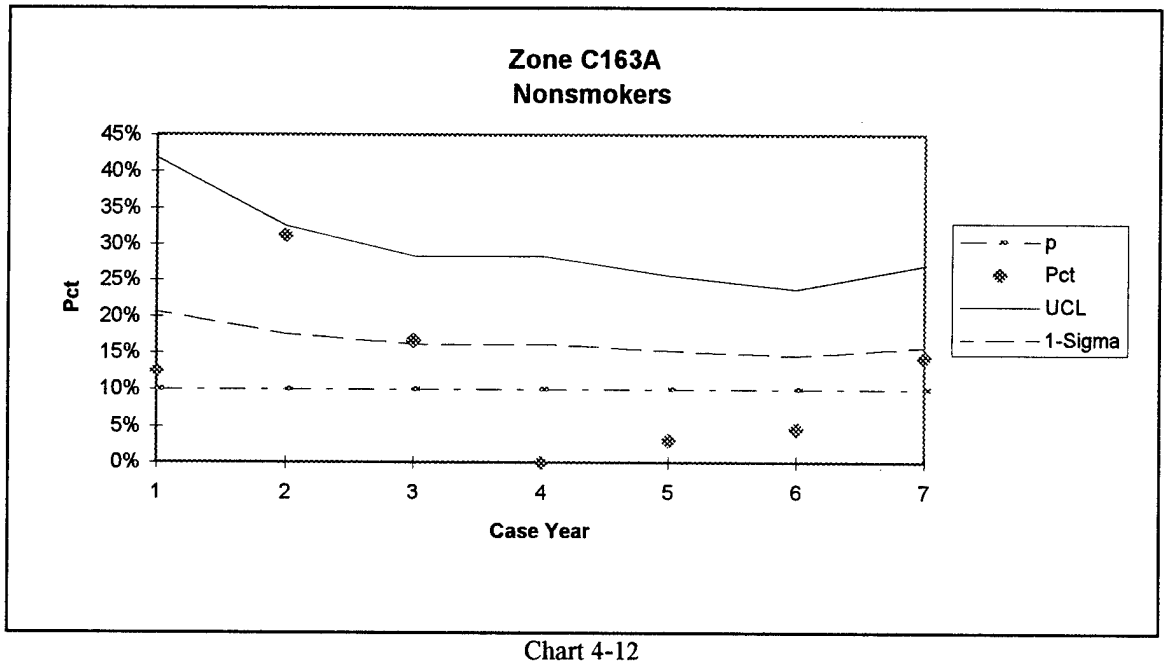


Chart 4-12
Zone C163A Nonsmokers

Zone C163A is out of control in Case years 2 and 7 with respect to it's entire population. However, when smokers are excluded, it is within the three standard deviation limit for Case year 2 and the one standard deviation limit for Case year 7.

Chapter 5 Summary and Recommendations

5.1 Summary of Key Results

The main result of this study is the 74th SGPMO now has a tool to use for analyzing the data collected on the WPAFB population. This tool is the *p*-chart for each respective work area . To maintain this analysis they only need to update the results from the respective work areas each year into a spreadsheet. The initial results of the study did show a large percentage of abnormalities base wide which indicates the population as a whole at WPAFB is unhealthy. Even with this large percentage of abnormalities, certain work areas over the years of the study displayed higher percentages of abnormal readings among their workers than the other work areas. These work areas were K1240, B36A, and B770A (all heat plants or heat distribution centers), C22I (the painters for the CE squadron), and C163A (firestations). The main recommendation from this study is for the 74th SGPMO to recognize these particular work areas as possible risk zones.

Section 5.2 The Smoker Issue

As stated above, the population monitored at WPAFB is on the whole unhealthy. Table 5-1 breaks down the total number of individual measurements used in the study by smoking status. Since both of these populations are binomial in nature (a test is either normal or abnormal), in comparing the two it is necessary to work with the proportion of abnormalities and test to see if this is the same (11:145). This test is a test of proportions and the logic of this test follows in Figure 5-1.

Table 5-1
Abnormalities by Smoking Status

Status	Total Measurements	Abnormal Results	Percent Abnormal
Overall	4544	801	18%
Nonsmokers	2090	287	14%
Smokers	2454	514	21%

- Null Hypothesis
  - $p_S = p_N$
- Alternative Hypothesis
  - $p_S > p_N$
- Test Statistic and $Z_{0.998}$
  - $Z^* = 6.34$; $Z_{0.998} = 3.09$
- Conclusion
  - Reject the Null

p_S is the percent of Smokers with an Abnormal Classification
p_N is the percent of Nonsmokers with an Abnormal Classification

Figure 5-1
Test of Proportions

The calculations for the above test can be found in appendix VII. Since we reject the Null Hypothesis, we conclude there is a true difference in the two populations. It is precisely for this reason that the nonsmoker population was analyzed on it's own and separate control charts were constructed for capturing possible hazardous work areas due to occupational exposures.

5.3 Recommendations and Ideas for Research

The first recommendation involves the initial data entry of the spirometry results into the computer. A system should be implemented to help reduce the number of bad data entries. The problem faced in this study was the majority of the ratio numbers were input as FVC or FEV_1. There was no way to logically check the validity of the FVC and FEV_1 entries once input into the database without going back to the initial paper record used to input the information. The current system has the individual accomplish the exam and a small printout of the results is generated by the computer; reference back to Figure 1-2. The information from this printout is then written on a form which is placed in the person's file. The information from this form is then read and input to the database by a data entry person. In each step of the process the potential for an error exists. Ideally, a Management Information System should be set up so the results of the exam are directly placed into the database from the spirometer itself; this will minimize human interaction and thus decrease the error rates.

The second recommendation involves the classification of a certain test as abnormal or normal. Instead of relying on a population predicted value which will classify at least 5% falsely as abnormal, add a box to the spirometry output in which the examiner can classify the test as properly on a "year-to-year" comparison basis or if the individual is tall and has a genetically reduced FEV_1 / FVC ratio. This individual's past test results must be accessible to the examiner for this to work. This result can then be entered into the database as a binary variable. This will reduce a great number of assumptions used in this paper in classifying a test as normal or abnormal. The main problem with respect to

the FEV_1 / FVC ratio is there exists no way for a programmer to determine if it is due to the individual being in great shape whereas the examiner would be in a much better position to make this determination.

The third recommendation is for a set of regression coefficients to be developed for the WPAFB population and to use these as normal values. This will help the examiner if the second recommendation is implemented or help the programmer feel more confident about the statistical classification of a test if this recommendation is not implemented.

The fourth recommendation is for the 74th SGPMO to conduct an experimental design for finding the true expected longitudinal decline based on age for each of the tests. Based on the results of such tests, a cutoff for abnormality can be implemented to aid the examiner in the second recommendation or the programmer in the third recommendation.

Appendix I: CONVERT.SAS Program

```
libname health 'user3';
run;
data health.demo;
 infile 'thesis1.';
 input first $ 1 ssan $ 1-9 yr 14-15 sex $ 17 race $ 21 ht 24-25 hyr 29-30 hmo 32-33 hdy 35-36;
 if index('0123456789',first)>0;
 format htdate yymmdd6.;
 htdate = mdy(hmo,hdy,hyr);
 drop hyr hmo hdy;
data health.lung;
 infile 'thesis2.';
 input first $ 1 ssan $ 1-9 lyr 13-14 lmo 16-17 ldy 19-20 fev1 22-25 fvc 27-30 ratio 32-35 fvcpred 38-40
fevpred 45-47 ratpred 52-54;
 if index('0123456789',first)>0;
 format lungdate yymmdd6.;
 lungdate = mdy(lmo,ldy,lyr);
 drop lmo ldy lyr;
data health.chem;
 infile 'thesis3.';
 input first $ 1 ssan $ 1-9 sgpt 11-13 sgot 16-18 yr 23-24 mo 26-27 dy 29-30  ap 32-33 ggt 37-39
 bili 42-45 albumin 52-55;
 if index('0123456789',first)>0;
 format chemdate yymmdd6.;
 chemdate = mdy(mo,dy,yr);
 drop mo dy yr;
data health.blood;
 infile 'thesis4.';
 input first $ 1 ssan $ 1-9 wbc $ 11-14 yr 20-21 mo 23-24 dy 26-27 hemcrit 29-33;
 if index('0123456789',first)>0;
 format blddate yymmdd6.;
 blddate = mdy(mo,dy,yr);
 drop mo dy yr;
data health.zone;
  infile 'thesis5.';
  input first $ 1 ssan $ 1-9 zone $ 11-18 syr 22-23 smo 25-26sdy 28-29 eyr 33-34 emo 36-37 edy 39-40;
 if index('0123456789',first)>0;
 format stdate yymmdd6. enddate yymmdd6.;
 stdate = mdy(smo,sdy,syr);
 enddate = mdy(emo,edy,eyr);
drop smo sdy syr emo edy eyr;
data health.history;
 infile 'thesis6.';
 input first $ 1 ssan $ 1-9 blddis $ 11 asthma $ 15 lung $ 19 liver $ 23 heptitis $ 27 jaundice $ 31 smk1 35
smk2 39-40   smk3 43-44;
 if index('0123456789',first)>0;
data health.drinking;
 infile 'thesis7.';
 input first $ 1 ssan $ 1-9 liquor 11-12 beer 15-16 wine 19-20;
 if index('0123456789',first)>0;
run;
```

Appendix II: FVCRAW.SAS Program

```
libname health 'user3';
run;
options ls = 75 ;
proc sort data = health.lung;
  by ssan;
run;
data _null_;
 set health.lung;
 by ssan;
 file print notitles;
if first.ssan then do;
 put @1 ssan @11 fvc  @;
 n = 16;
end;
if first.ssan = 0 and last.ssan = 0 then do;
 put @n fvc  @;
 n = n+5;
 retain n;
end;
if last.ssan then do;
 put @n fvc  @75  first;
end;
run;
```

Appendix III: MERGEALL.SAS Program

```
data a;
 infile 'demo.raw';
 input ssan $ 1-9 dobyr 11-12 sex $ 14 race $ 15
ht 16-17;
run;
proc sort data = a; by ssan ; run;

data b;
 infile 'fev1.raw';
 input ssan $ 1-9 fev11 11-15 fev12 16-20 fev13
21-25 fev14 26-30
     fev15 31-35 fev16 36-40 fev17 41-45 fev18
46-50;
run;
proc sort data = b; by ssan ; run;

data c;
 infile 'fvc.raw';
 input ssan $ 1-9 fvc1  11-15 fvc2  16-20 fvc3
21-25 fvc4  26-30
     fvc5  31-35 fvc6  36-40 fvc7  41-45 fvc8
46-50;
run;
proc sort data = c; by ssan ; run;

data d;
 infile 'ratio.raw';
 input ssan $ 1-9 ratio1 11-15 ratio2 16-20
ratio3 21-25 ratio4 26-30
    ratio5 31-35 ratio6 36-40 ratio7 41-45 ratio8
46-50;
run;
proc sort data = d; by ssan ; run;

data e;
 infile 'fvcpred.raw';
 input ssan $ 1-9 fvcpd1 11-15 fvcpd2 16-20
fvcpd3 21-25 fvcpd4 26-30
    fvcpd5 31-35 fvcpd6 36-40 fvcpd7 41-45
fvcpd8 46-50;
run;
proc sort data = e; by ssan ; run;

data f;
 infile 'fevpred.raw';
 input ssan $ 1-9 fevpd1 11-15 fevpd2 16-20
fevpd3 21-25 fevpd4 26-30
    fevpd5 31-35 fevpd6 36-40 fevpd7 41-45
fevpd8 46-50;
run;
proc sort data = f; by ssan ; run;

data g;
 infile 'ratpred.raw';
 input ssan $ 1-9 ratpd1 11-15 ratpd2 16-20
ratpd3 21-25 ratpd4 26-30
    ratpd5 31-35 ratpd6 36-40 ratpd7 41-45
ratpd8 46-50;
run;
proc sort data = g; by ssan ; run;

data h;
 infile 'lungdate.raw';
 input ssan $ 1-9 yr1 11-12 mo1 13-14 dy1 15-
16 yr2 18-19 mo2 20-21
     dy2 22-23 yr3 25-26 mo3 27-28 dy3 29-30
yr4 32-33 mo4 34-35
     dy4 36-37 yr5 39-40 mo5 41-42 dy5 43-44
yr6 46-47 mo6 48-49
     dy6 50-51 yr7 53-54 mo7 54-56 dy7 57-58
yr8 60-61 mo8 62-63
     dy8 64-65;
 format ldt1 ldt2 ldt3 ldt4 ldt5 ldt6 ldt7 ldt8
yymmdd6.;
 ldt1 = mdy(mo1,dy1,yr1);
 ldt2 = mdy(mo2,dy2,yr2);
 ldt3 = mdy(mo3,dy3,yr3);
 ldt4 = mdy(mo4,dy4,yr4);
 ldt5 = mdy(mo5,dy5,yr5);
 ldt6 = mdy(mo6,dy6,yr6);
 ldt7 = mdy(mo7,dy7,yr7);
 ldt8 = mdy(mo8,dy8,yr8);
drop yr1 mo1 dy1 yr2 mo2 dy2 yr3 mo3 dy3 yr4
mo4 dy4 yr5 mo5 dy5 yr6
   mo6 dy6 yr7 mo7 dy7 yr8 mo8 dy8;
run;
proc sort data = h; by ssan ; run;

data i;
 infile 'sgpt.raw';
 input ssan $ 1-9 sgpt1  11-14 sgpt2  15-18
sgpt3  19-22 sgpt4  23-26
    sgpt5  27-30 sgpt6  31-34 sgpt7  35-38 sgpt8
39-42;
run;
proc sort data = i; by ssan ; run;

data j;
```

```
 infile 'sgot.raw';
 input ssan $ 1-9 sgot1  11-14 sgot2  15-18 sgot3
19-22 sgot4  23-26
    sgot5  27-30 sgot6  31-34 sgot7  35-38 sgot8
39-42;
run;
proc sort data = j; by ssan ; run;

data k;
 infile 'ap.raw'  ;
 input ssan $ 1-9 ap1    11-14 ap2    15-18 ap3
19-22 ap4    23-26
    ap5    27-30 ap6    31-34 ap7    35-38 ap8
39-42;
run;
proc sort data = k; by ssan ; run;

data l;
 infile 'ggt.raw' ;
 input ssan $ 1-9 ggt1   11-14 ggt2   15-18 ggt3
19-22 ggt4   23-26
    ggt5   27-30 ggt6   31-34 ggt7   35-38 ggt8
39-42;
run;
proc sort data = l; by ssan ; run;


data m;
 infile 'bili.raw';
 input ssan $ 1-9 bili1  11-15 bili2  16-20 bili3
21-25 bili4  26-30
    bili5  31-35 bili6  36-40 bili7  41-45 bili8
46-51;
run;
proc sort data = m; by ssan ; run;

data n;
 infile 'albumin.raw' ;
 input ssan $ 1-9 albumin1 11-15 albumin2 16-
20 albumin3 21-25
    albumin4 26-30 albumin5 31-35 albumin6
36-40 albumin7 41-45
    albumin8 46-50;
run;
proc sort data = n; by ssan ; run;

data o;
 infile 'chemdate.raw';
 input ssan $ 1-9 yr1 11-12 mo1 13-14 dy1 15-
16 yr2 18-19 mo2 20-21
    dy2 22-23 yr3 26-27 mo3 28-29 dy3 30-31
yr4 34-35 mo4 36-37
    dy4 38-39 yr5 42-43 mo5 44-45 dy5 46-47
yr6 50-51 mo6 52-53
    dy6 54-55 yr7 58-59 mo7 60-61 dy7 62-63
yr8 66-67 mo8 68-69
    dy8 70-71;
 format cdt1 cdt2 cdt3 cdt4 cdt5 cdt6 cdt7 cdt8
yymmdd6.;
 cdt1 = mdy(mo1,dy1,yr1);
 cdt2 = mdy(mo2,dy2,yr2);
 cdt3 = mdy(mo3,dy3,yr3);
 cdt4 = mdy(mo4,dy4,yr4);
 cdt5 = mdy(mo5,dy5,yr5);
 cdt6 = mdy(mo6,dy6,yr6);
 cdt7 = mdy(mo7,dy7,yr7);
 cdt8 = mdy(mo8,dy8,yr8);
drop yr1 mo1 dy1 yr2 mo2 dy2 yr3 mo3 dy3 yr4
mo4 dy4 yr5 mo5 dy5 yr6
   mo6 dy6 yr7 mo7 dy7 yr8 mo8 dy8;
run;
proc sort data = o; by ssan ; run;

data p;
 infile 'blddate.raw' ;
 input ssan $ 1-9 yr1 11-12 mo1 13-14 dy1 15-
16 yr2 18-19 mo2 20-21
    dy2 22-23 yr3 26-27 mo3 28-29 dy3 30-31
yr4 34-35 mo4 36-37
    dy4 38-39 yr5 42-43 mo5 44-45 dy5 46-47
yr6 50-51 mo6 52-53
    dy6 54-55 yr7 58-59 mo7 60-61 dy7 62-63
yr8 66-67 mo8 68-69
    dy8 70-71;
 format bdt1 bdt2 bdt3 bdt4 bdt5 bdt6 bdt7 bdt8
yymmdd6.;
 bdt1 = mdy(mo1,dy1,yr1);
 bdt2 = mdy(mo2,dy2,yr2);
 bdt3 = mdy(mo3,dy3,yr3);
 bdt4 = mdy(mo4,dy4,yr4);
 bdt5 = mdy(mo5,dy5,yr5);
 bdt6 = mdy(mo6,dy6,yr6);
 bdt7 = mdy(mo7,dy7,yr7);
 bdt8 = mdy(mo8,dy8,yr8);
drop yr1 mo1 dy1 yr2 mo2 dy2 yr3 mo3 dy3 yr4
mo4 dy4 yr5 mo5 dy5 yr6
   mo6 dy6 yr7 mo7 dy7 yr8 mo8 dy8;
run;
proc sort data = p; by ssan ; run;

data q;
 infile 'wbc.raw' ;
 input ssan $ 1-9 wbc1   11-17 wbc2   18-25
wbc3   26-33 wbc4   34-41
```

```
    wbc5  42-49 wbc6  50-57 wbc7  58-65
wbc8  66-74;
run;
proc sort data = q; by ssan ; run;

data r;
 infile 'hemcrit.raw';
 input ssan $ 1-9 hem1  11-17 hem2  18-25
hem3  26-33 hem4  34-41
    hem5  42-49 hem6  50-57 hem7  58-65
hem8  66-74;
run;
proc sort data = r; by ssan ; run;

data s;
 infile 'zone.raw'  ;
 input ssan $ 1-9 zone1 $11-19 zone2 $20-28
zone3 $29-37 zone4 $38-46
    zone5 $47-55 zone6 $56-64 zone7 $65-73 ;
run;
proc sort data = s; by ssan ; run;

data t;
 infile 'stdate.raw'  ;
 input ssan $ 1-9 yr1 11-12 mo1 13-14 dy1 15-
16 yr2 20-21 mo2 22-23
     dy2 24-25 yr3 29-30 mo3 31-32 dy3 33-34
yr4 38-39 mo4 40-41
     dy4 42-43 yr5 47-48 mo5 49-50 dy5 51-52
yr6 56-57 mo6 58-59
     dy6 60-61 yr7 65-66 mo7 67-68 dy7 69-70;
 format sdt1 sdt2 sdt3 sdt4 sdt5 sdt6 sdt7
yymmdd6.;
 sdt1 = mdy(mo1,dy1,yr1);
 sdt2 = mdy(mo2,dy2,yr2);
 sdt3 = mdy(mo3,dy3,yr3);
 sdt4 = mdy(mo4,dy4,yr4);
 sdt5 = mdy(mo5,dy5,yr5);
 sdt6 = mdy(mo6,dy6,yr6);
 sdt7 = mdy(mo7,dy7,yr7);
drop yr1 mo1 dy1 yr2 mo2 dy2 yr3 mo3 dy3 yr4
mo4 dy4 yr5 mo5 dy5 yr6
   mo6 dy6 yr7 mo7 dy7;
run;
proc sort data = t; by ssan ; run;

data u;
 infile 'enddate.raw' ;
 input ssan $ 1-9 yr1 11-12 mo1 13-14 dy1 15-
16 yr2 20-21 mo2 22-23
     dy2 24-25 yr3 29-30 mo3 31-32 dy3 33-34
yr4 38-39 mo4 40-41
     dy4 42-43 yr5 47-48 mo5 49-50 dy5 51-52
yr6 56-57 mo6 58-59
     dy6 60-61 yr7 65-66 mo7 67-68 dy7 69-70;
 format edt1 edt2 edt3 edt4 edt5 edt6 edt7
yymmdd6.;
 edt1 = mdy(mo1,dy1,yr1);
 edt2 = mdy(mo2,dy2,yr2);
 edt3 = mdy(mo3,dy3,yr3);
 edt4 = mdy(mo4,dy4,yr4);
 edt5 = mdy(mo5,dy5,yr5);
 edt6 = mdy(mo6,dy6,yr6);
 edt7 = mdy(mo7,dy7,yr7);
drop yr1 mo1 dy1 yr2 mo2 dy2 yr3 mo3 dy3 yr4
mo4 dy4 yr5 mo5 dy5 yr6
   mo6 dy6 yr7 mo7 dy7;
run;
proc sort data = u; by ssan ; run;

data v;
 infile 'history.raw';
 input ssan $ 1-9 blddis $ 11 asthma $ 12 lung $
13 liver $ 14
     heptitis $ 15 jaundice $ 16 smk1 17-18
smk2 19-21 smk3 22-23;
run;
proc sort data = v; by ssan ; run;

data w;
 infile 'drinking.raw';
 input ssan $ 1-9 liquor 11-13 beer 14-16 wine
17-19;
run;
proc sort data = w; by ssan ; run;


libname health 'user3';
data health.wpafb1;
  merge a b (in=in1) c d e f g h i (in=in2) j k l m
n o p q (in=in3)
      r s t u v w;
 by ssan;
 if in1 or in2 or in3;
 if zone1 = '  ' then delete;
run;
proc contents;
run;
```

Appendix IV: LUNGYRS1.SAS and LUNGYRS2.SAS Programs

Section IV.1 LUNGYRS1.SAS Program

```
libname health 'user3';
run;
data a y1987 y1988 y1989 y1990 y1991 y1992
y1993 y1994 y1995;
 set health.wpafb2 (keep = ldt1-ldt8  fvc1-fvc8
fev11-fev18 ssan
    ratio1-ratio8 fvcpd1-fvcpd8 fevpd1-fevpd8
ratpd1-ratpd8
    zone1-zone7 sdt1-sdt7 edt1-edt7);
 array  fvc[*]  fvc1-fvc8;
 array fev1[*] fev11-fev18;
 array ratio[*] ratio1-ratio8;
 array ratpd[*] ratpd1-ratpd8;
 array fvcpd[*] fvcpd1-fvcpd8;
 array fevpd[*] fevpd1-fevpd8;
 array ldt[*] ldt1-ldt8;
 array ldtyr[*] ldtyr1-ldtyr8;
 array keep[*] keep1-keep8;
 array zone[*] zone1-zone8;
 array sdt[*] sdt1-sdt8;
 array edt[*] edt1-edt8;
 zone8 = '       ';
 sdt8 = .;
 edt8 = .;
 format cratio 6.2;
/*
  Create Proper Work Areas
*/
  do j = 1 to 7;
   if  zone[j] = 'A1400' then  zone[j] =
substr(zone[j],1,6);
   else if  zone[j] = 'B6000' then  zone[j] =
substr(zone[j],1,6);
   else if  zone[j] = 'C4020' then  zone[j] =
substr(zone[j],1,6);
   else if  zone[j] = 'K1084' then  zone[j] =
substr(zone[j],1,6);
 else if
index('B70.B18.B22.B36.B63.B79.C89',substr(
zone[j],1,3))> 0
     then   zone[j] = substr(zone[j],1,4);
 else if
index('B40.C13.C19.C22.C28.C59.C70',substr(
zone[j],1,3))> 0
     then   zone[j] = substr(zone[j],1,4);
   else if substr(zone[j],1,4) = 'B24A' then
zone[j] = 'B24A';
   else if substr(zone[j],1,4) = 'B24B' then
zone[j] = 'B24B';
   else if substr(zone[j],1,4) = 'B45B' then
zone[j] = 'B45B';
   else if substr(zone[j],1,4) = 'B65A' then
zone[j] = 'B65A';
   else if substr(zone[j],1,3) = 'B4D' then zone[j]
= 'B4D';
   else if substr(zone[j],1,3) = 'B4E' then zone[j]
= 'B4E';
   else if substr(zone[j],1,3) = 'B5C' then zone[j]
= 'B5C';
   else if substr(zone[j],1,3) = 'B5D' then zone[j]
= 'B5D';
   else if substr(zone[j],1,3) = 'B5E' then zone[j]
= 'B5E';
   else if substr(zone[j],1,3) = 'B5H' then zone[j]
= 'B5H';
   else if substr(zone[j],1,2) = 'CE' then zone[j]
= 'CE';
   else if substr(zone[j],1,3) = 'C40'
     then zone[j] = substr(zone[j],1,6);
   else zone[j] = substr(zone[j],1,5);
end;
do j = 1 to 7;
 if sdt[j] ^= . and edt[j] = . then edt[j] =
'31dec99'd;
end;
 do j = 1 to 8;
  if  fvc[j] ^= . then keep[j] = 1;
 end;
 if keep1 = 1 or keep2 = 1 or keep3 = 1 or keep4
= 1 or keep5 = 1
   or keep6 = 1 or keep7 = 1 or keep8 = 1;
 do j = 1 to 8;
  if keep[j] = . then keep[j] = 0;
  if keep1 = keep2 = 1 and ldt1 = . then ldt1 =
ldt2-365;
 end;
 do j = 1 to 7;
  if keep[j] = keep[j+1] = 1 then do;
    if ldt[j] = ldt[j+1] then do;
      ldt[j+1] = .;
      keep[j+1] = 0;
    end;
  end;
 end;
do j = 1 to 8;
 if keep[j] = 1 then ldtyr[j] = year(ldt[j]);
 if ldtyr[j] ^= . then do;
  if ldtyr[j] < 1988 then do;
```

```
        vc =  fvc[j];
        fev = fev1[j];
     fev_fvc = ratio[j];
     predfvc = fvcpd[j];
     predfev = fevpd[j];
     prdratio= ratpd[j];
       tyear = ldtyr[j];
     if vc ^= . and fev ^= . then cratio
=(fev/vc)*100;
     if sdt[j] <= ldt[j] <= edt[j] then czone =
zone[j];
     if substr(czone,1,1) = ' ' then czone = zone1;
        output y1987;
     end;
     else if ldtyr[j] = 1988 then do;
         vc =  fvc[j];
        fev = fev1[j];
     fev_fvc = ratio[j];
     predfvc = fvcpd[j];
     predfev = fevpd[j];
     prdratio= ratpd[j];
       tyear = ldtyr[j];
     if vc ^= . and fev ^= . then cratio
=(fev/vc)*100;
     if sdt[j] <= ldt[j] <= edt[j] then czone =
zone[j];
     if substr(czone,1,1) = ' ' then czone = zone1;
        output y1988;
     end;
     else if ldtyr[j] = 1989 then do;
         vc =  fvc[j];
        fev = fev1[j];
     fev_fvc = ratio[j];
     predfvc = fvcpd[j];
     predfev = fevpd[j];
     prdratio= ratpd[j];
       tyear = ldtyr[j];
     if vc ^= . and fev ^= . then cratio
=(fev/vc)*100;
     if sdt[j] <= ldt[j] <= edt[j] then czone =
zone[j];
     if substr(czone,1,1) = ' ' then czone = zone1;
        output y1989;
     end;
     else if ldtyr[j] = 1990 then do;
         vc =  fvc[j];
        fev = fev1[j];
     fev_fvc = ratio[j];
     predfvc = fvcpd[j];
     predfev = fevpd[j];
     prdratio= ratpd[j];
       tyear = ldtyr[j];
     if vc ^= . and fev ^= . then cratio
=(fev/vc)*100;
     if sdt[j] <= ldt[j] <= edt[j] then czone =
zone[j];
     if substr(czone,1,1) = ' ' then czone = zone1;
        output y1990;
     end;
     else if ldtyr[j] = 1991 then do;
         vc =  fvc[j];
        fev = fev1[j];
     fev_fvc = ratio[j];
     predfvc = fvcpd[j];
     predfev = fevpd[j];
     prdratio= ratpd[j];
       tyear = ldtyr[j];
     if vc ^= . and fev ^= . then cratio
=(fev/vc)*100;
     if sdt[j] <= ldt[j] <= edt[j] then czone =
zone[j];
     if substr(czone,1,1) = ' ' then czone = zone1;
        output y1991;
     end;
     else if ldtyr[j] = 1992 then do;
         vc =  fvc[j];
        fev = fev1[j];
     fev_fvc = ratio[j];
     predfvc = fvcpd[j];
     predfev = fevpd[j];
     prdratio= ratpd[j];
       tyear = ldtyr[j];
     if vc ^= . and fev ^= . then cratio
=(fev/vc)*100;
     if sdt[j] <= ldt[j] <= edt[j] then czone =
zone[j];
     if substr(czone,1,1) = ' ' then czone = zone1;
        output y1992;
     end;
     else if ldtyr[j] = 1993 then do;
         vc =  fvc[j];
        fev = fev1[j];
     fev_fvc = ratio[j];
     predfvc = fvcpd[j];
     predfev = fevpd[j];
     prdratio= ratpd[j];
       tyear = ldtyr[j];
     if vc ^= . and fev ^= . then cratio
=(fev/vc)*100;
     if sdt[j] <= ldt[j] <= edt[j] then czone =
zone[j];
     if substr(czone,1,1) = ' ' then czone = zone1;
        output y1993;
     end;
```

```
    else if ldtyr[j] = 1994 then do;
        vc =  fvc[j];
       fev = fev1[j];
    fev_fvc = ratio[j];
    predfvc = fvcpd[j];
    predfev = fevpd[j];
    prdratio= ratpd[j];
      tyear = ldtyr[j];
    if vc ^= . and fev ^= . then cratio
=(fev/vc)*100;
    if sdt[j] <= ldt[j] <= edt[j] then czone =
zone[j];
    if substr(czone,1,1) = ' ' then czone = zone1;
       output y1994;
    end;
    else if ldtyr[j] = 1995 then do;
        vc =  fvc[j];
       fev = fev1[j];
    fev_fvc = ratio[j];
    predfvc = fvcpd[j];
    predfev = fevpd[j];
    prdratio= ratpd[j];
      tyear = ldtyr[j];
    if vc ^= . and fev ^= . then cratio
=(fev/vc)*100;
    if sdt[j] <= ldt[j] <= edt[j] then czone =
zone[j];
    if substr(czone,1,1) = ' ' then czone = zone1;
       output y1995;
    end;
  end;
end;
run;
data y1987;
 set y1987 (keep = ssan       vc fev fev_fvc
predfvc predfev prdratio
              cratio czone);
 vc87 = vc; fev87 = fev; ratio87 = fev_fvc;
pfvc87 = predfvc;
 pfev87 = predfev; pratio87 = prdratio; cratio87
= cratio;
 czone87 = czone;
 id = ssan||'_87';
run;
data y1988;
 set y1988 (keep = ssan  vc fev fev_fvc predfvc
predfev prdratio
              cratio czone);
 vc88 = vc; fev88 = fev; ratio88 = fev_fvc;
pfvc88 = predfvc;
 pfev88 = predfev; pratio88 = prdratio; cratio88
= cratio;
 czone88 = czone;
 id = ssan||'_88';
run;
data y1989;
 set y1989 (keep = ssan       vc fev fev_fvc
predfvc predfev prdratio
              cratio czone);
 vc89 = vc; fev89 = fev; ratio89 = fev_fvc;
pfvc89 = predfvc;
 pfev89 = predfev; pratio89 = prdratio; cratio89
= cratio;
 czone89 = czone;
 id = ssan||'_89';
run;
data y1990;
 set y1990 (keep = ssan  vc fev fev_fvc predfvc
predfev prdratio
              cratio czone);
 vc90 = vc; fev90 = fev; ratio90 = fev_fvc;
pfvc90 = predfvc;
 pfev90 = predfev; pratio90 = prdratio; cratio90
= cratio;
 czone90 = czone;
 id = ssan||'_90';
run;
data y1991;
 set y1991 (keep = ssan       vc fev fev_fvc
predfvc predfev prdratio
              cratio czone);
 vc91 = vc; fev91 = fev; ratio91 = fev_fvc;
pfvc91 = predfvc;
 pfev91 = predfev; pratio91 = prdratio; cratio91
= cratio;
 czone91 = czone;
 id = ssan||'_91';
run;
data y1992;
 set y1992 (keep = ssan vc fev fev_fvc predfvc
predfev prdratio
              cratio czone);
 vc92 = vc; fev92 = fev; ratio92 = fev_fvc;
pfvc92 = predfvc;
 pfev92 = predfev; pratio92 = prdratio; cratio92
= cratio;
 czone92 = czone;
 id = ssan||'_92';
run;
data y1993;
 set y1993 (keep = ssan       vc fev fev_fvc
predfvc predfev prdratio
              cratio czone);
```

```
 vc93 = vc; fev93 = fev; ratio93 = fev_fvc;
pfvc93 = predfvc;
 pfev93 = predfev; pratio93 = prdratio; cratio93
= cratio;
 czone93 = czone;
 id = ssan||'_93';
run;
data y1994;
 set y1994 (keep = ssan vc fev fev_fvc predfvc
predfev prdratio
                cratio czone);
 vc94 = vc; fev94 = fev; ratio94 = fev_fvc;
pfvc94 = predfvc;
 pfev94 = predfev; pratio94 = prdratio; cratio94
= cratio;
 czone94 = czone;
 id = ssan||'_94';
run;
data y1995;
 set y1995 (keep = ssan       vc fev fev_fvc
predfvc predfev prdratio
                cratio czone);
 vc95 = vc; fev95 = fev; ratio95 = fev_fvc;
pfvc95 = predfvc;
 pfev95 = predfev; pratio95 = prdratio; cratio95
= cratio;
 czone95 = czone;
 id = ssan||'_95';
run;
data health.lung (drop = vc fev fev_fvc predfvc
predfev prdratio cratio
                    czone);
 merge y1987 y1988 y1989 y1990 y1991 y1992
y1993 y1994 y1995;
 by ssan;
run;
proc contents data = health.lung;
run;
```

Section IV.2 LUNGYRS2.SAS Program

```
libname health 'user3';
run;
data a;
 set health.lung3;
 array cratio[*] cratio87-cratio95;
 array czone[*] czone87-czone95;
 array fev[*] fev87-fev95;
 array pfev[*] pfev87-pfev95;
 array pfvc[*] pfvc87-pfvc95;
 array pratio[*] pratio87-pratio95;
 array ratio[*] ratio87-ratio95;
 array abfvc[*] abfvc1-abfvc9;
 array abfev[*] abfev1-abfev9;
 array abrat[*] abrat1-abrat9;
 array ablfvc[*] ablfvc1-ablfvc9;
 array ablfev[*] ablfev1-ablfev9;
 array ablrat[*] ablrat1-ablrat9;
 array vc[*] vc87-vc95;
 array keep[*] keep1-keep9;
 array pval[*] pval1-pval9;
 array pvala[*] pvala1-pvala9;
 array smoke[*] smoke1-smoke9;
 do j =  1 to 9 ;
  if vc[j] ^= . then keep[j] = 1; else keep[j] = 0;
  if keep[j] = 1 then smoke[j] = smoker; else
smoke[j] = 0;
 end;
/*
 Define Abnormalities based on Percent
Predicted Values of FVC, FEV1
 and the Observed Value of the Ratio
FEV1/FVC
*/

if dobyr > . and ht > . then do;
  htcm = 2.54*ht;
 do j = 1 to 9;
  age = (j+86)-dobyr;
 if sex = 'M' then do;
  if age < 25 then do;
   pval[j] = -4.808 + .045*age + .046*htcm;
  end;
  else do;
   pval[j] = -4.203 - .027*age + .052*htcm;
  end;
 end;

 else if sex = 'F' then do;
  if age < 25 then do;
   pval[j] = -2.703 + .085*age + .027*htcm;
  end;
  else do;
   pval[j] = -0.794 - .021*age + .027*htcm;
  end;
 end;
 pfev[j] = (fev[j]/pval[j])*100;
end;

do j = 1 to 9;
  age = (j+86)-dobyr;
 if sex = 'M' then do;
```

```
    if age < 25 then do;
     pvala[j] = -5.508 + .078*age + .05*htcm;
    end;
    else do;
     pvala[j] = -5.459 - .029*age + .065*htcm;
    end;
   end;

  else if sex = 'F'  then do;
    if age < 20 then do;
     pvala[j] = -3.469 + .092*age + .033*htcm;
    end;
    else do;
     pvala[j] = -1.774 - .022*age + .037*htcm;
    end;
   end;
  pfvc[j] = (vc[j]/pvala[j])*100;
end;
end;
 if dobyr = . then dobyr = 50;
 do j = 1 to 9;
 if keep[j] = 1 then do;
 if sex = 'M' or sex = ' ' then do;
   if ((86+j) - dobyr) > 39 then do;
    if pfvc[j] ^= . then do;
     if pfvc[j] < 73.4 then abfvc[j] = 1; else
abfvc[j] = 0;
    end;
    else abfvc[j] = 0;
    if pfev[j] ^= . then do;
     if pfev[j] < 77.2 then abfev[j] = 1; else
abfev[j] = 0;
    end;
    else abfev[j] = 0;
   end;
   else do;
    if pfvc[j] ^= . then do;
     if pfvc[j] < 81.1 then abfvc[j] = 1; else
abfvc[j] = 0;
    end;
    else abfvc[j] = 0;
    if pfev[j] ^= . then do;
     if pfev[j] < 79.1 then abfev[j] = 1; else
abfev[j] = 0;
    end;
    else abfev[j] = 0;
  end;
 end;
else if sex = 'F' then do;
  if ((86+j) - dobyr) > 39 then do;
   if pfvc[j] ^= . then do;
     if pfvc[j] < 75.2 then abfvc[j] = 1; else
abfvc[j] = 0;
    end;
    else abfvc[j] = 0;
    if pfev[j] ^= . then do;
     if pfev[j] < 77.9 then abfev[j] = 1; else
abfev[j] = 0;
    end;
    else abfev[j] = 0;
   end;
   else do;
    if pfvc[j] ^= . then do;
     if pfvc[j] < 76.9 then abfvc[j] = 1; else
abfvc[j] = 0;
    end;
    else abfvc[j] = 0;
    if pfev[j] ^= . then do;
     if pfev[j] < 70.3 then abfev[j] = 1; else
abfev[j] = 0;
    end;
    else abfev[j] = 0;
  end;
 end;
  if cratio[j] ^= . then do;
   if cratio[j] < 70 then abrat[j] = 1; else abrat[j]
= 0;
  end;
  else abrat[j] = 0;
end;
end;
/*
 Define abnormalities based on longitudinal
data.  Based on
 Percentage decrease of observed FVC, FEV1,
and Ratio
*/
do j = 1 to 8;
  if keep[j] = keep[j+1] = 1 then do;
   if vc[j] > 0 and vc[j+1] > 0 and vc[j] > vc[j+1]
then do;
    if ((vc[j]-vc[j+1])/vc[j]) > .15 then
ablfvc[j+1] = 1;
    else ablfvc[j+1] = 0;
  end;
  if fev[j] > 0 and fev[j+1] > 0 and fev[j] >
fev[j+1] then do;
    if ((fev[j]-fev[j+1])/fev[j]) > .15 then
ablfev[j+1] = 1;
    else ablfev[j+1] = 0;
  end;
  if cratio[j] > 0 and cratio[j+1] > 0 and cratio[j]
> cratio[j+1]
```

```
   then do;
    if ((cratio[j]-cratio[j+1])/cratio[j]) > .10 then
ablrat[j+1] = 1;
    else ablrat[j+1] = 0;
   end;
  end;
 else do;
   ablfvc[j+1] = 0;
   ablfev[j+1] = 0;
   ablrat[j+1] = 0;
end;
end;
do j = 1;
   ablfvc[j] = 0;
   ablfev[j] = 0;
   ablrat[j] = 0;
end;
/*
do j = 1 to 7;
  if keep[j] = keep[j+2] = 1 then do;
   if vc[j] > 0 and vc[j+2] > 0 and vc[j] > vc[j+2]
then do;
    if ((vc[j]-vc[j+2])/vc[j]) > .15 then
ablfvc[j+2] = 1;
   end;
   if fev[j] > 0 and fev[j+2] > 0 and fev[j] >
fev[j+2] then do;
    if ((fev[j]-fev[j+2])/fev[j]) > .15 then
ablfev[j+2] = 1;
   end;
   if cratio[j] > 0 and cratio[j+2] > 0 and cratio[j]
> cratio[j+2]
    then do;
    if ((cratio[j]-cratio[j+2])/cratio[j]) > .05 then
ablrat[j+2] = 1;
   end;
  end;
end;

do j = 1 to 6;
  if keep[j] = keep[j+3] = 1 then do;
   if vc[j] > 0 and vc[j+3] > 0 and vc[j] > vc[j+3]
then do;
    if ((vc[j]-vc[j+3])/vc[j]) > .15 then
ablfvc[j+3] = 1;
   end;
   if fev[j] > 0 and fev[j+3] > 0 and fev[j] >
fev[j+3] then do;
    if ((fev[j]-fev[j+3])/fev[j]) > .15 then
ablfev[j+3] = 1;
   end;
   if cratio[j] > 0 and cratio[j+3] > 0 and cratio[j]
> cratio[j+3]
    then do;
    if ((cratio[j]-cratio[j+3])/cratio[j]) > .05 then
ablrat[j+3] = 1;
   end;
  end;
end;

do j = 1 to 5;
  if keep[j] = keep[j+4] = 1 then do;
   if vc[j] > 0 and vc[j+4] > 0 and vc[j] > vc[j+4]
then do;
    if ((vc[j]-vc[j+4])/vc[j]) > .15 then
ablfvc[j+4] = 1;
   end;
   if fev[j] > 0 and fev[j+4] > 0 and fev[j] >
fev[j+4] then do;
    if ((fev[j]-fev[j+4])/fev[j]) > .15 then
ablfev[j+4] = 1;
   end;
   if cratio[j] > 0 and cratio[j+4] > 0 and cratio[j]
> cratio[j+4]
    then do;
    if ((cratio[j]-cratio[j+4])/cratio[j]) > .05 then
ablrat[j+4] = 1;
   end;
  end;
end;

do j = 1 to 4;
  if keep[j] = keep[j+5] = 1 then do;
   if vc[j] > 0 and vc[j+5] > 0 and vc[j] > vc[j+5]
then do;
    if ((vc[j]-vc[j+5])/vc[j]) > .15 then
ablfvc[j+5] = 1;
   end;
   if fev[j] > 0 and fev[j+5] > 0 and fev[j] >
fev[j+5] then do;
    if ((fev[j]-fev[j+5])/fev[j]) > .15 then
ablfev[j+5] = 1;
   end;
   if cratio[j] > 0 and cratio[j+5] > 0 and cratio[j]
> cratio[j+5]
    then do;
    if ((cratio[j]-cratio[j+5])/cratio[j]) > .05 then
ablrat[j+5] = 1;
   end;
  end;
end;

do j = 1 to 3;
```

```
  if keep[j] = keep[j+6] = 1 then do;
   if vc[j] > 0 and vc[j+6] > 0 and vc[j] > vc[j+6]
then do;
    if ((vc[j]-vc[j+6])/vc[j]) > .15 then
ablfvc[j+6] = 1;
  end;
  if fev[j] > 0 and fev[j+6] > 0 and fev[j] >
fev[j+6] then do;
    if ((fev[j]-fev[j+6])/fev[j]) > .15 then
ablfev[j+6] = 1;
  end;
  if cratio[j] > 0 and cratio[j+6] > 0 and cratio[j]
> cratio[j+6]
   then do;
    if ((cratio[j]-cratio[j+6])/cratio[j]) > .05 then
ablrat[j+6] = 1;
  end;
 end;
end;

do j = 1 to 2;
 if keep[j] = keep[j+7] = 1 then do;
  if vc[j] > 0 and vc[j+7] > 0 and vc[j] > vc[j+7]
then do;
    if ((vc[j]-vc[j+7])/vc[j]) > .15 then
ablfvc[j+7] = 1;
  end;
  if fev[j] > 0 and fev[j+7] > 0 and fev[j] >
fev[j+7] then do;
    if ((fev[j]-fev[j+7])/fev[j]) > .15 then
ablfev[j+7] = 1;
  end;
  if cratio[j] > 0 and cratio[j+7] > 0 and cratio[j]
> cratio[j+7]
   then do;
    if ((cratio[j]-cratio[j+7])/cratio[j]) > .05 then
ablrat[j+7] = 1;
  end;
 end;
end;
*/
do j = 1 to 9;
 if abfvc[j] = . then abfvc[j] = 0;
 if abfev[j] = . then abfev[j] = 0;
 if abrat[j] = . then abrat[j] = 0;
 if ablfvc[j] = . then ablfvc[j] = 0;
 if ablfev[j] = . then ablfev[j] = 0;
 if ablrat[j] = . then ablrat[j] = 0;
end;
run;

data y1987;
 set a;
 array abfvc[*] abfvc1-abfvc9;
 array abfev[*] abfev1-abfev9;
 array abrat[*] abrat1-abrat9;
 array ablfvc[*] ablfvc1-ablfvc9;
 array ablfev[*] ablfev1-ablfev9;
 array ablrat[*] ablrat1-ablrat9;
 array vc[*] vc87-vc95;
 array czone[*] czone87-czone95;
 array keep[*] keep1-keep9;
 array ab[*] ab1-ab9;
 array sab[*] sab1-sab9;
 array smoke[*] smoke1-smoke9;
 j = 1;
 if keep[j] = 1;
 zone = czone87;
 count1 = 1;
 ab[j] =
abfvc[j]+abfev[j]+abrat[j]+ablfvc[j]+ablfev[j]+ab
lrat[j];
 if ab[j] > 0 then ab[j] = 1;
 else ab[j] = 0;
 if ab[j] = 1 and smoke[j] = 1 then sab[j] = 1;
  else sab[j] = 0;
run;
proc summary;
class zone;
var smoke1 count1 abfvc1 abfev1 abrat1 ablfvc1
ablfev1 ablrat1 ab1 sab1;
 output out = y87 sum = ;
run;

data y1988;
 set a;
 array abfvc[*] abfvc1-abfvc9;
 array abfev[*] abfev1-abfev9;
 array abrat[*] abrat1-abrat9;
 array ablfvc[*] ablfvc1-ablfvc9;
 array ablfev[*] ablfev1-ablfev9;
 array ablrat[*] ablrat1-ablrat9;
 array vc[*] vc87-vc95;
 array czone[*] czone87-czone95;
 array ab[*] ab1-ab9;
 array sab[*] sab1-sab9;
 array smoke[*] smoke1-smoke9;
 array keep[*] keep1-keep9;
 j = 2;
 if keep[j] = 1;
 zone = czone88;
 count2 = 1;
```

```
 ab[j] =
abfvc[j]+abfev[j]+abrat[j]+ablfvc[j]+ablfev[j]+ab
lrat[j];
 if ab[j] > 0 then ab[j] = 1;
 else ab[j] = 0;
 if ab[j] = 1 and smoke[j] = 1 then sab[j] = 1;
  else sab[j] = 0;
run;
proc summary;
class zone;
var smoke2 count2 abfvc2 abfev2 abrat2 ablfvc2
ablfev2 ablrat2 ab2 sab2;
 output out = y88 sum = ;
run;

data y1989;
 set a;
 array abfvc[*] abfvc1-abfvc9;
 array abfev[*] abfev1-abfev9;
 array abrat[*] abrat1-abrat9;
 array ablfvc[*] ablfvc1-ablfvc9;
 array ablfev[*] ablfev1-ablfev9;
 array ablrat[*] ablrat1-ablrat9;
 array vc[*] vc87-vc95;
 array czone[*] czone87-czone95;
 array ab[*] ab1-ab9;
 array sab[*] sab1-sab9;
 array smoke[*] smoke1-smoke9;
 array keep[*] keep1-keep9;
 j = 3;
 if keep[j] = 1;
 zone = czone89;
 count3 = 1;
 ab[j] =
abfvc[j]+abfev[j]+abrat[j]+ablfvc[j]+ablfev[j]+ab
lrat[j];
 if ab[j] > 0 then ab[j] = 1;
 else ab[j] = 0;
 if ab[j] = 1 and smoke[j] = 1 then sab[j] = 1;
  else sab[j] = 0;
run;
proc summary;
class zone;
var smoke3 count3 abfvc3 abfev3 abrat3 ablfvc3
ablfev3 ablrat3 ab3 sab3;
 output out = y89 sum = ;
run;

data y1990;
 set a;
 array abfvc[*] abfvc1-abfvc9;
 array abfev[*] abfev1-abfev9;
 array abrat[*] abrat1-abrat9;
 array ablfvc[*] ablfvc1-ablfvc9;
 array ablfev[*] ablfev1-ablfev9;
 array ablrat[*] ablrat1-ablrat9;
 array vc[*] vc87-vc95;
 array czone[*] czone87-czone95;
 array ab[*] ab1-ab9;
 array sab[*] sab1-sab9;
 array smoke[*] smoke1-smoke9;
 array keep[*] keep1-keep9;
 j = 4;
 if keep[j] = 1;
 zone = czone90;
 count4 = 1;
 ab[j] =
abfvc[j]+abfev[j]+abrat[j]+ablfvc[j]+ablfev[j]+ab
lrat[j];
 if ab[j] > 0 then ab[j] = 1;
 else ab[j] = 0;
 if ab[j] = 1 and smoke[j] = 1 then sab[j] = 1;
  else sab[j] = 0;
run;
proc summary;
class zone;
var smoke4 count4 abfvc4 abfev4 abrat4 ablfvc4
ablfev4 ablrat4 ab4 sab4;
 output out = y90 sum = ;
run;

data y1991;
 set a;
 array abfvc[*] abfvc1-abfvc9;
 array abfev[*] abfev1-abfev9;
 array abrat[*] abrat1-abrat9;
 array ablfvc[*] ablfvc1-ablfvc9;
 array ablfev[*] ablfev1-ablfev9;
 array ablrat[*] ablrat1-ablrat9;
 array vc[*] vc87-vc95;
 array czone[*] czone87-czone95;
 array ab[*] ab1-ab9;
 array sab[*] sab1-sab9;
 array smoke[*] smoke1-smoke9;
 array keep[*] keep1-keep9;
 j = 5;
 if keep[j] = 1;
 zone = czone91;
 count5 = 1;
 ab[j] =
abfvc[j]+abfev[j]+abrat[j]+ablfvc[j]+ablfev[j]+ab
lrat[j];
 if ab[j] > 0 then ab[j] = 1;
 else ab[j] = 0;
```

```
 if ab[j] = 1 and smoke[j] = 1 then sab[j] = 1;
  else sab[j] = 0;
run;
proc summary;
class zone;
var smoke5 count5 abfvc5 abfev5 abrat5 ablfvc5
ablfev5 ablrat5 ab5 sab5;
 output out = y91 sum = ;
run;

data y1992;
 set a;
 array abfvc[*] abfvc1-abfvc9;
 array abfev[*] abfev1-abfev9;
 array abrat[*] abrat1-abrat9;
 array ablfvc[*] ablfvc1-ablfvc9;
 array ablfev[*] ablfev1-ablfev9;
 array ablrat[*] ablrat1-ablrat9;
 array vc[*] vc87-vc95;
 array czone[*] czone87-czone95;
 array ab[*] ab1-ab9;
 array sab[*] sab1-sab9;
 array smoke[*] smoke1-smoke9;
 array keep[*] keep1-keep9;
 j = 6;
 if keep[j] = 1;
 zone = czone92;
 count6 = 1;
 ab[j] =
abfvc[j]+abfev[j]+abrat[j]+ablfvc[j]+ablfev[j]+ab
lrat[j];
 if ab[j] > 0 then ab[j] = 1;
 else ab[j] = 0;
 if ab[j] = 1 and smoke[j] = 1 then sab[j] = 1;
  else sab[j] = 0;
run;
proc summary;
class zone;
var smoke6 count6 abfvc6 abfev6 abrat6 ablfvc6
ablfev6 ablrat6 ab6 sab6;
 output out = y92 sum = ;
run;

data y1993;
 set a;
 array abfvc[*] abfvc1-abfvc9;
 array abfev[*] abfev1-abfev9;
 array abrat[*] abrat1-abrat9;
 array ablfvc[*] ablfvc1-ablfvc9;
 array ablfev[*] ablfev1-ablfev9;
 array ablrat[*] ablrat1-ablrat9;
 array vc[*] vc87-vc95;
 array ab[*] ab1-ab9;
 array sab[*] sab1-sab9;
 array smoke[*] smoke1-smoke9;
 array czone[*] czone87-czone95;
 array keep[*] keep1-keep9;
 j = 7;
 if keep[j] = 1;
 zone = czone93;
 count7 = 1;
 ab[j] =
abfvc[j]+abfev[j]+abrat[j]+ablfvc[j]+ablfev[j]+ab
lrat[j];
 if ab[j] > 0 then ab[j] = 1;
 else ab[j] = 0;
 if ab[j] = 1 and smoke[j] = 1 then sab[j] = 1;
  else sab[j] = 0;
run;
proc summary;
class zone;
var smoke7 count7 abfvc7 abfev7 abrat7 ablfvc7
ablfev7 ablrat7 ab7 sab7;
 output out = y93 sum = ;
run;

data y1994;
 set a;
 array abfvc[*] abfvc1-abfvc9;
 array abfev[*] abfev1-abfev9;
 array abrat[*] abrat1-abrat9;
 array ablfvc[*] ablfvc1-ablfvc9;
 array ablfev[*] ablfev1-ablfev9;
 array ablrat[*] ablrat1-ablrat9;
 array vc[*] vc87-vc95;
 array czone[*] czone87-czone95;
 array ab[*] ab1-ab9;
 array sab[*] sab1-sab9;
 array smoke[*] smoke1-smoke9;
 array keep[*] keep1-keep9;
 j = 8;
 if keep[j] = 1;
 zone = czone94;
 count8 = 1;
 ab[j] =
abfvc[j]+abfev[j]+abrat[j]+ablfvc[j]+ablfev[j]+ab
lrat[j];
 if ab[j] > 0 then ab[j] = 1;
 else ab[j] = 0;
 if ab[j] = 1 and smoke[j] = 1 then sab[j] = 1;
  else sab[j] = 0;
run;
proc summary;
class zone;
```

```
var smoke8 count8 abfvc8 abfev8 abrat8 ablfvc8
ablfev8 ablrat8 ab8 sab8;
 output out = y94 sum = ;
run;

data y1995;
 set a;
 array abfvc[*] abfvc1-abfvc9;
 array abfev[*] abfev1-abfev9;
 array abrat[*] abrat1-abrat9;
 array ablfvc[*] ablfvc1-ablfvc9;
 array ablfev[*] ablfev1-ablfev9;
 array ablrat[*] ablrat1-ablrat9;
 array vc[*] vc87-vc95;
 array czone[*] czone87-czone95;
 array ab[*] ab1-ab9;
 array sab[*] sab1-sab9;
 array smoke[*] smoke1-smoke9;
 array keep[*] keep1-keep9;
 j = 9;
 if keep[j] = 1;
 zone = czone95;
 count9 = 1;
 ab[j] =
abfvc[j]+abfev[j]+abrat[j]+ablfvc[j]+ablfev[j]+ab
lrat[j];
 if ab[j] > 0 then ab[j] = 1;
 else ab[j] = 0;
 if ab[j] = 1 and smoke[j] = 1 then sab[j] = 1;
  else sab[j] = 0;
run;
proc summary;
class zone;
var smoke9 count9 abfvc9 abfev9 abrat9 ablfvc9
ablfev9 ablrat9 ab9 sab9;
 output out = y95 sum = ;
run;

data d1 d2 d3 d4 d5 d6 d7 d8 d9 f;
 merge y87 y88 y89 y90 y91 y92 y93 y94 y95;
 by zone;
 array count[*] count1-count9;
 array abfvc[*] abfvc1-abfvc9;
 array abfev[*] abfev1-abfev9;
 array abrat[*] abrat1-abrat9;
 array ablfvc[*] ablfvc1-ablfvc9;
 array ablfev[*] ablfev1-ablfev9;
 array ablrat[*] ablrat1-ablrat9;
 array d[*] d1-d9;
 array ab[*] ab1-ab9;
 array sab[*] sab1-sab9;
 array smoke[*] smoke1-smoke9;
 do j = 1 to 9;
 if j = 1 and count1 > 4 then output d1;
 else if j = 2 and count2 > 4 then output d2;
 else if j = 3 and count3 > 4 then output d3;
 else if j = 4 and count4 > 4 then output d4;
 else if j = 5 and count5 > 4 then output d5;
 else if j = 6 and count6 > 4 then output d6;
 else if j = 7 and count7 > 4 then output d7;
 else if j = 8 and count8 > 4 then output d8;
 else if j = 9 and count9 > 4 then output d9;
 end;
run;
data l1;
 set d1;
 format pct 6.2 pctsmk 6.2;
 pct = (ab1 / count1) * 100;
 pctsmk = (smoke1 / count1)*100;
run;
proc sort; by zone; run;
proc print;
 var zone count1 ab1 pct smoke1 pctsmk sab1;
 title '1987 zones';
run;

data l2;
 set d2;
 format pct 6.2 pctsmk 6.2;
 pct = (ab2 / count2) * 100;
 pctsmk = (smoke2 / count2)*100;
run;
proc sort; by zone; run;
proc print;
 var zone count2 ab2 pct smoke2 pctsmk sab2;
 title '1988 zones';
run;

data l3;
 set d3;
 format pct 6.2 pctsmk 6.2;
 pct = (ab3 / count3) * 100;
 pctsmk = (smoke3 / count3)*100;
run;
proc sort; by zone; run;
proc print;
 var zone count3 ab3 pct smoke3 pctsmk sab3;
 title '1989 zones';
run;

data l4;
 set d4;
 format pct 6.2 pctsmk 6.2;
 pct = (ab4 / count4) * 100;
```

```
 pctsmk = (smoke4 / count4)*100;
run;
proc sort; by zone; run;
proc print;
 var zone count4 ab4 pct smoke4 pctsmk sab4;
 title '1990 zones';
run;

data l5;
 set d5;
 format pct 6.2 pctsmk 6.2;
 pct = (ab5 / count5) * 100;
 pctsmk = (smoke5 / count5)*100;
run;
proc sort; by zone; run;
proc print;
 var zone count5 ab5 pct smoke5 pctsmk sab5;
 title '1991 zones';
run;

data l6;
 set d6;
 format pct 6.2 pctsmk 6.2;
 pct = (ab6 / count6) * 100;
 pctsmk = (smoke6 / count6)*100;
run;
proc sort; by zone; run;
proc print;
 var zone count6 ab6 pct smoke6 pctsmk sab6;
 title '1992 zones';
run;

data l7;
 set d7;
 format pct 6.2 pctsmk 6.2;
 pct = (ab7 / count7) * 100;
 pctsmk = (smoke7 / count7) * 100;
run;
proc sort; by zone; run;
proc print;
 var zone count7 ab7 pct smoke7 pctsmk sab7;
 title '1993 zones';
run;

data l8;
 set d8;
 format pct 6.2 pctsmk 6.2;
 pct = (ab8 / count8) * 100;
 pctsmk = (smoke8 / count8)*100;
run;
proc sort; by zone; run;
proc print;
 var zone count8 ab8 pct smoke8 pctsmk sab8;
 title '1994 zones';
run;

data l9;
 set d9;
 format pct 6.2 pctsmk 6.2;
 pct = (ab9 / count9) * 100;
 pctsmk = (smoke9 / count9) * 100;
run;
proc sort; by zone; run;
proc print;
 var zone count9 ab9 pct smoke9 pctsmk sab9;
 title '1995 zones';
run;
```

Appendix V: Work Area Description and Case Number Identifiers

Section V.1: Work Area Description

ZONE	Description
A1400	88th CCSG, SCTWO Unit : Telephone Installation and Repair
A1405	88th CCSG, SCLR Unit: Metereological Center
A262E	AFMC Headquarters Building
A278A	88th CEG, CEMIUE: Pest Management
A830A-Q	Hospital, each letter identifies unique exposure areas
A862A	Unknown
A867A	NAIC, LGMM: Operations Maintenance
A876C	88th CEG, Grounds Area and Pavement Equipment Area
A878A	Military Golf Course
B18G	Wright Labs: Experimental Research Branch
B248A	Armstrong Labs: Hazard Assessment
B24A	Wright Labs: Aero-Diagnostics Research
B254A	Wright Labs: Signature Technology
B36A-F	88th CEG, Heat Distribution Area B
B40A	74th SGB, Bioenvironmental Engineering
B433A	NMRI, Navy Toxicology
B470B	AFIT, ENP unit: Nuclear Spectrum
B490A,D	Wright Labs: Experimental Support Branch and Avionics Facility
B4B1,D	Wright Labs: Electro-Optics Warfare
B5B1,C	ASC: Machine A/B and Machine Repair
B6000	Firestations
B620F	Wright Labs: Microwave Division
B640B	AFIT / ENP unit: Physics
B652D	Wright Labs: Material and Surface Interaction
B654A,B	Wright Labs: Mechanics Interactions and Non-Structural Materials
B65A	Wright Labs: Fat Frac/Rel Grp
B682A	645th LCMPS: Library of Congress and Motion Pictures
B70A	SA-ALC: Aerospace Fuels Lab
B745B,C,D	88th CEG, Pavement Equipment, Water Sewer & Gas, Heat Plant Area B
B76A1,A2	88th CEG, Firestations
B770A	88th CEG, Heat Plant Area B
B79A,B,C,E	Armstrong Labs: Hazard Assessment
B838A	Armstrong Labs: Occupational Environmental Vet Medicine
C101B,I	Unknown
C105A	Unknown
C106C	Unknown
C13D,F,O,P,R	445th LG, Machine Welding, Propulsion, Survival Equipment, Wheel / Tire
C163A	88th CEG, Fire Stations
C170A	88th CEG, Heat Plant Area C
C174A	74th Med Grp, Orthopedic Brace Shop
C19A,C	88th CEG, Hazardous Material & Waste, Water Treatment
C206A,C,D,E	ASC, Aircraft Modification Division

ZONE	Description
C21C1	Unknown
C22A,B,C,I	88th CEG, Cathodic Protection, CE zone A,B, Project Painters
C255A	Unknown
C28B	Unknown
C29A1	Unknown
C4014A	Unknown
C4020A,B	44th LGMAF, Fuel Systems
C4021E	88th ABW, Age Section
C4024A	Unknown
C4035A	88th ABW, Survival Equipment Repair
C4066A	Unknown
C4067A	Unknown
C59A	71st Ordnance Attachment: Explosive Ordnance
C70B	AFOSI: AFOSI Technical Systems
C883A	88th SPS: Combat Arms Training
C884A	88th ABW: PMEL
C89B	DRMO Area C
C91B1	Unknown
CE	CE Worker Areas
K1084	88th ABW, Frame Shop / Wood Hobby Shop
K1240	Heat Plant Kittyhawk Area

Section V.2 Speadsheet Calculations for Chart 4-1

Case #	ZONE	Year	Total	Tot Abnm	Pct	p	Zi
1	A1400	89	13	5	0.38	0.1	3.420651
2	A1400	90	44	12	0.27	0.1	3.819144
3	A1405	90	8	0	0.00	0.1	-0.94281
4	A1405	91	22	2	0.09	0.1	-0.14213
5	A1405	92	17	0	0.00	0.1	-1.37437
6	A1405	93	18	0	0.00	0.1	-1.41421
7	A1405	94	16	0	0.00	0.1	-1.33333
8	A262E	90	16	1	0.06	0.1	-0.5
9	A262E	91	8	1	0.13	0.1	0.235702
10	A278A	89	6	1	0.17	0.1	0.544331
11	A278A	90	17	6	0.35	0.1	3.476344
12	A278A	91	13	0	0.00	0.1	-1.20185
13	A278A	92	14	2	0.14	0.1	0.534522
14	A278A	93	13	4	0.31	0.1	2.496151
15	A278A	94	14	1	0.07	0.1	-0.35635
16	A830A	90	9	1	0.11	0.1	0.111111
17	A830A	91	15	2	0.13	0.1	0.430331
18	A830A	92	10	1	0.10	0.1	0
19	A830A	93	19	1	0.05	0.1	-0.68825
20	A830A	94	11	1	0.09	0.1	-0.1005

Case #	ZONE	Year	Total	Tot Abnm	Pct	p	Zi
21	A830A	95	27	3	0.11	0.1	0.19245
22	A830C	90	10	0	0.00	0.1	-1.05409
23	A830C	91	8	0	0.00	0.1	-0.94281
24	A830C	92	13	0	0.00	0.1	-1.20185
25	A830C	93	24	1	0.04	0.1	-0.95258
26	A830C	94	18	1	0.06	0.1	-0.62854
27	A830D	93	10	0	0.00	0.1	-1.05409
28	A830F	89	17	9	0.53	0.1	5.9017
29	A830F	90	31	8	0.26	0.1	2.933553
30	A830F	93	6	3	0.50	0.1	3.265986
31	A830Q	92	5	0	0.00	0.1	-0.74536
32	A830Q	93	6	1	0.17	0.1	0.544331
33	A830Q	94	5	1	0.20	0.1	0.745356
34	A862A	90	5	1	0.20	0.1	0.745356
35	A862A	92	6	0	0.00	0.1	-0.8165
36	A862A	93	6	0	0.00	0.1	-0.8165
37	A867A	90	7	2	0.29	0.1	1.637846
38	A867A	92	6	1	0.17	0.1	0.544331
39	A867A	93	6	3	0.50	0.1	3.265986
40	A876C	89	5	2	0.40	0.1	2.236068
41	A876C	90	21	9	0.43	0.1	5.019011
42	A876C	92	19	4	0.21	0.1	1.60591
43	A876C	93	27	4	0.15	0.1	0.83395
44	A876C	95	9	1	0.11	0.1	0.111111
45	A878A	90	6	1	0.17	0.1	0.544331
46	A878A	91	5	1	0.20	0.1	0.745356
47	A878A	92	5	1	0.20	0.1	0.745356
48	A878A	93	7	1	0.14	0.1	0.377964
49	ALL	89	232	59	0.25	0.1	7.834617
50	ALL	90	923	215	0.23	0.1	13.4624
51	ALL	91	854	137	0.16	0.1	5.885719
52	ALL	92	733	105	0.14	0.1	3.902887
53	ALL	93	836	127	0.15	0.1	5.003401
54	ALL	94	676	104	0.15	0.1	4.666667
55	ALL	95	290	54	0.19	0.1	4.893502
56	B18G	94	6	1	0.17	0.1	0.544331
57	B248A	91	8	1	0.13	0.1	0.235702
58	B24A	91	8	2	0.25	0.1	1.414214
59	B24A	92	8	0	0.00	0.1	-0.94281
60	B254A	90	6	0	0.00	0.1	-0.8165
61	B36A	89	21	4	0.19	0.1	1.382047
62	B36A	90	35	13	0.37	0.1	5.352644
63	B36A	91	16	4	0.25	0.1	2
64	B36A	92	20	5	0.25	0.1	2.236068
65	B36A	93	13	2	0.15	0.1	0.64715
66	B36A	94	16	6	0.38	0.1	3.666667

Case #	ZONE	Year	Total	Tot Abnm	Pct	p	Zi
67	B36A	95	8	3	0.38	0.1	2.592725
68	B36B	90	15	3	0.20	0.1	1.290994
69	B36B	91	8	2	0.25	0.1	1.414214
70	B36C	90	5	1	0.20	0.1	0.745356
71	B36F	91	12	1	0.08	0.1	-0.19245
72	B40A	90	21	2	0.10	0.1	-0.07274
73	B40A	91	26	3	0.12	0.1	0.261488
74	B40A	92	31	2	0.06	0.1	-0.65855
75	B40A	93	34	4	0.12	0.1	0.342997
76	B40A	94	5	0	0.00	0.1	-0.74536
77	B40A	95	5	0	0.00	0.1	-0.74536
78	B433A	90	6	0	0.00	0.1	-0.8165
79	B433A	91	20	1	0.05	0.1	-0.74536
80	B433A	92	9	0	0.00	0.1	-1
81	B433A	93	36	1	0.03	0.1	-1.44444
82	B433A	94	29	1	0.03	0.1	-1.17607
83	B470B	91	10	5	0.50	0.1	4.21637
84	B470B	92	10	2	0.20	0.1	1.054093
85	B470B	93	10	1	0.10	0.1	0
86	B470B	94	5	0	0.00	0.1	-0.74536
87	B490A	90	13	5	0.38	0.1	3.420651
88	B490A	91	11	6	0.55	0.1	4.924685
89	B490D	93	7	2	0.29	0.1	1.637846
90	B490D	94	5	1	0.20	0.1	0.745356
91	B4B1	90	5	1	0.20	0.1	0.745356
92	B4B1	91	9	4	0.44	0.1	3.444444
93	B4B1	92	16	8	0.50	0.1	5.333333
94	B4B1	93	16	2	0.13	0.1	0.333333
95	B4B1	94	14	1	0.07	0.1	-0.35635
96	B4B1	95	9	1	0.11	0.1	0.111111
97	B4D	89	16	4	0.25	0.1	2
98	B4D	90	24	11	0.46	0.1	5.851559
99	B4D	91	19	2	0.11	0.1	0.076472
100	B4D	92	15	1	0.07	0.1	-0.43033
101	B4D	93	30	7	0.23	0.1	2.434322
102	B4D	94	12	3	0.25	0.1	1.732051
103	B5B1	90	6	2	0.33	0.1	1.905159
104	B5B1	91	7	1	0.14	0.1	0.377964
105	B5C	90	15	2	0.13	0.1	0.430331
106	B5C	93	11	0	0.00	0.1	-1.10554
107	B5C	94	9	0	0.00	0.1	-1
108	B5E	92	5	3	0.60	0.1	3.72678
109	B5H	92	6	0	0.00	0.1	-0.8165
110	B5J1	90	22	1	0.05	0.1	-0.8528
111	B5J1	91	26	3	0.12	0.1	0.261488
112	B5J1	92	23	3	0.13	0.1	0.486534

Case #	ZONE	Year	Total	Tot Abnm	Pct	p	Zi
113	B5J1	93	24	2	0.08	0.1	-0.27217
114	B5J1	94	21	1	0.05	0.1	-0.80013
115	B5J1	95	18	6	0.33	0.1	3.299832
116	B6000	90	26	3	0.12	0.1	0.261488
117	B6000	91	19	4	0.21	0.1	1.60591
118	B6000	92	13	4	0.31	0.1	2.496151
119	B6000	93	7	1	0.14	0.1	0.377964
120	B6000	94	6	1	0.17	0.1	0.544331
121	B620F	91	20	2	0.10	0.1	0
122	B640B	90	5	0	0.00	0.1	-0.74536
123	B640B	91	5	1	0.20	0.1	0.745356
124	B652D	94	5	0	0.00	0.1	-0.74536
125	B654A	90	6	0	0.00	0.1	-0.8165
126	B654B	91	5	2	0.40	0.1	2.236068
127	B654B	92	6	0	0.00	0.1	-0.8165
128	B654B	93	6	1	0.17	0.1	0.544331
129	B65A	90	14	6	0.43	0.1	4.098006
130	B65A	91	19	5	0.26	0.1	2.370629
131	B65A	92	16	1	0.06	0.1	-0.5
132	B65A	94	6	1	0.17	0.1	0.544331
133	B682A	89	7	1	0.14	0.1	0.377964
134	B682A	90	9	2	0.22	0.1	1.222222
135	B682A	91	10	2	0.20	0.1	1.054093
136	B682A	92	10	2	0.20	0.1	1.054093
137	B682A	93	11	2	0.18	0.1	0.904534
138	B682A	94	13	0	0.00	0.1	-1.20185
139	B70A	91	13	2	0.15	0.1	0.64715
140	B745B	90	6	0	0.00	0.1	-0.8165
141	B745B	91	8	4	0.50	0.1	3.771236
142	B745B	92	6	2	0.33	0.1	1.905159
143	B745C	94	7	1	0.14	0.1	0.377964
144	B745D	92	5	1	0.20	0.1	0.745356
145	B745D	93	13	3	0.23	0.1	1.571651
146	B745D	94	14	4	0.29	0.1	2.316264
147	B745D	95	7	0	0.00	0.1	-0.88192
148	B76A1	89	9	2	0.22	0.1	1.222222
149	B76A1	90	36	5	0.14	0.1	0.777778
150	B76A1	91	30	4	0.13	0.1	0.608581
151	B76A1	92	18	3	0.17	0.1	0.942809
152	B76A1	93	14	3	0.21	0.1	1.425393
153	B76A1	94	14	1	0.07	0.1	-0.35635
154	B76A1	95	12	2	0.17	0.1	0.7698
155	B76A2	91	5	2	0.40	0.1	2.236068
156	B770A	89	20	6	0.30	0.1	2.981424
157	B770A	90	38	17	0.45	0.1	7.137743
158	B770A	91	34	4	0.12	0.1	0.342997

Case #	ZONE	Year	Total	Tot Abnm	Pct	p	Zi
159	B770A	92	37	9	0.24	0.1	2.904382
160	B770A	93	35	6	0.17	0.1	1.40859
161	B770A	94	33	7	0.21	0.1	2.146958
162	B770A	95	21	4	0.19	0.1	1.382047
163	B79A	90	22	0	0.00	0.1	-1.56347
164	B79A	91	34	1	0.03	0.1	-1.37199
165	B79A	92	24	1	0.04	0.1	-0.95258
166	B79A	93	22	0	0.00	0.1	-1.56347
167	B79A	94	16	2	0.13	0.1	0.333333
168	B79B	91	8	1	0.13	0.1	0.235702
169	B79B	92	8	1	0.13	0.1	0.235702
170	B79B	93	10	0	0.00	0.1	-1.05409
171	B79B	94	5	0	0.00	0.1	-0.74536
172	B79C	91	6	0	0.00	0.1	-0.8165
173	B79E	94	17	1	0.06	0.1	-0.56592
174	B838A	90	6	0	0.00	0.1	-0.8165
175	B838A	94	5	0	0.00	0.1	-0.74536
176	C101B	92	5	0	0.00	0.1	-0.74536
177	C101I	89	5	0	0.00	0.1	-0.74536
178	C101I	90	5	0	0.00	0.1	-0.74536
179	C101I	94	5	0	0.00	0.1	-0.74536
180	C105A	90	7	2	0.29	0.1	1.637846
181	C106C	92	10	2	0.20	0.1	1.054093
182	C13D	91	7	2	0.29	0.1	1.637846
183	C13D	92	9	2	0.22	0.1	1.222222
184	C13D	93	7	4	0.57	0.1	4.157609
185	C13F	89	5	0	0.00	0.1	-0.74536
186	C13F	90	6	3	0.50	0.1	3.265986
187	C13O	94	5	0	0.00	0.1	-0.74536
188	C13P	90	5	2	0.40	0.1	2.236068
189	C13R	89	7	0	0.00	0.1	-0.88192
190	C13R	91	9	3	0.33	0.1	2.333333
191	C13R	92	18	3	0.17	0.1	0.942809
192	C13R	93	20	3	0.15	0.1	0.745356
193	C13R	94	10	4	0.40	0.1	3.162278
194	C163A	89	13	2	0.15	0.1	0.64715
195	C163A	90	38	12	0.32	0.1	4.434052
196	C163A	91	50	8	0.16	0.1	1.414214
197	C163A	92	57	4	0.07	0.1	-0.75057
198	C163A	93	73	8	0.11	0.1	0.273096
199	C163A	94	90	10	0.11	0.1	0.351364
200	C163A	95	65	14	0.22	0.1	3.100868
201	C170A	90	25	3	0.12	0.1	0.333333
202	C170A	91	8	0	0.00	0.1	-0.94281
203	C174A	92	5	0	0.00	0.1	-0.74536
204	C19A	90	18	4	0.22	0.1	1.728483

Case #	ZONE	Year	Total	Tot Abnm	Pct	p	Zi
205	C19A	91	11	3	0.27	0.1	1.909572
206	C19C	94	6	0	0.00	0.1	-0.8165
207	C206A	93	10	0	0.00	0.1	-1.05409
208	C206A	94	8	0	0.00	0.1	-0.94281
209	C206C	93	9	1	0.11	0.1	0.111111
210	C206D	90	7	0	0.00	0.1	-0.88192
211	C206D	93	9	1	0.11	0.1	0.111111
212	C206E	90	9	0	0.00	0.1	-1
213	C206E	91	12	1	0.08	0.1	-0.19245
214	C206E	92	7	0	0.00	0.1	-0.88192
215	C206E	93	5	0	0.00	0.1	-0.74536
216	C21C1	91	14	2	0.14	0.1	0.534522
217	C22A	90	15	5	0.33	0.1	3.01232
218	C22A	91	14	3	0.21	0.1	1.425393
219	C22A	92	8	1	0.13	0.1	0.235702
220	C22A	93	14	3	0.21	0.1	1.425393
221	C22A	94	13	3	0.23	0.1	1.571651
222	C22A	95	6	1	0.17	0.1	0.544331
223	C22B	90	16	4	0.25	0.1	2
224	C22B	91	15	2	0.13	0.1	0.430331
225	C22C	90	11	1	0.09	0.1	-0.1005
226	C22I	90	10	3	0.30	0.1	2.108185
227	C22I	91	12	3	0.25	0.1	1.732051
228	C22I	92	11	2	0.18	0.1	0.904534
229	C22I	93	14	6	0.43	0.1	4.098006
230	C22I	94	7	2	0.29	0.1	1.637846
231	C255A	90	15	3	0.20	0.1	1.290994
232	C28B	90	5	1	0.20	0.1	0.745356
233	C28B	91	7	0	0.00	0.1	-0.88192
234	C28B	92	7	1	0.14	0.1	0.377964
235	C28B	93	7	1	0.14	0.1	0.377964
236	C28B	94	7	1	0.14	0.1	0.377964
237	C29A1	89	7	4	0.57	0.1	4.157609
238	C29A1	90	13	1	0.08	0.1	-0.27735
239	C29A1	91	13	2	0.15	0.1	0.64715
240	C29A1	92	12	1	0.08	0.1	-0.19245
241	C29A1	93	11	1	0.09	0.1	-0.1005
242	C29A1	94	10	4	0.40	0.1	3.162278
243	C4014A	90	6	2	0.33	0.1	1.905159
244	C4020A	89	5	2	0.40	0.1	2.236068
245	C4020A	90	7	4	0.57	0.1	4.157609
246	C4020A	91	7	0	0.00	0.1	-0.88192
247	C4020A	92	9	1	0.11	0.1	0.111111
248	C4020A	93	8	1	0.13	0.1	0.235702
249	C4020B	94	5	0	0.00	0.1	-0.74536
250	C4020B	95	5	0	0.00	0.1	-0.74536

Case #	ZONE	Year	Total	Tot Abnm	Pct	p	Zi
251	C4021E	93	6	0	0.00	0.1	-0.8165
252	C4021E	94	5	1	0.20	0.1	0.745356
253	C4024A	90	5	0	0.00	0.1	-0.74536
254	C4035A	91	9	2	0.22	0.1	1.222222
255	C4035A	92	5	2	0.40	0.1	2.236068
256	C4035A	93	6	2	0.33	0.1	1.905159
257	C4066A	91	9	1	0.11	0.1	0.111111
258	C4066A	92	8	0	0.00	0.1	-0.94281
259	C4066A	93	14	1	0.07	0.1	-0.35635
260	C4066A	94	12	4	0.33	0.1	2.694301
261	C4067A	92	6	0	0.00	0.1	-0.8165
262	C4067A	93	6	2	0.33	0.1	1.905159
263	C59A	93	6	0	0.00	0.1	-0.8165
264	C59A	94	11	0	0.00	0.1	-1.10554
265	C59A	95	11	0	0.00	0.1	-1.10554
266	C70B	90	10	2	0.20	0.1	1.054093
267	C70B	91	8	2	0.25	0.1	1.414214
268	C70B	92	7	1	0.14	0.1	0.377964
269	C883A	94	5	3	0.60	0.1	3.72678
270	C884A	91	6	1	0.17	0.1	0.544331
271	C89B	90	5	0	0.00	0.1	-0.74536
272	C89B	91	7	0	0.00	0.1	-0.88192
273	C89B	92	11	0	0.00	0.1	-1.10554
274	C89B	94	24	1	0.04	0.1	-0.95258
275	C91B1	92	5	2	0.40	0.1	2.236068
276	C91B1	93	5	0	0.00	0.1	-0.74536
277	C91B1	94	5	3	0.60	0.1	3.72678
278	CE	95	8	1	0.13	0.1	0.235702
279	K1084	93	6	2	0.33	0.1	1.905159
280	K1240	89	29	12	0.41	0.1	5.632759
281	K1240	90	33	17	0.52	0.1	7.949546
282	K1240	91	28	6	0.21	0.1	2.015811
283	K1240	92	30	13	0.43	0.1	6.085806
284	K1240	93	28	9	0.32	0.1	3.905633
285	K1240	94	28	9	0.32	0.1	3.905633
286	K1240	95	22	5	0.23	0.1	1.989873

Appendix VI: Critical Zone Calculations

Zone K1240: Total Population

Year	Total	Abnormal	p	Pct	UCL	1-Sigma	Case Year
89	29	12	10%	41%	27%	16%	1
90	33	17	10%	52%	26%	15%	2
91	28	6	10%	21%	27%	16%	3
92	30	13	10%	43%	26%	15%	4
93	28	9	10%	32%	27%	16%	5
94	28	9	10%	32%	27%	16%	6
95	22	5	10%	23%	29%	16%	7

Zone K1240: Nonsmoker Only Population

Year	Total	Abnormal	p	Pct	UCL	1-Sigma	Case Year
89	8	4	0.1	50%	42%	21%	1
90	9	4	0.1	44%	40%	20%	2
91	9	2	0.1	22%	40%	20%	3
92	9	5	0.1	56%	40%	20%	4
93	8	3	0.1	38%	42%	21%	5
94	9	4	0.1	44%	40%	20%	6
95	8	3	0.1	38%	42%	21%	7

Zone B36A: Total Population

Year	Total	Tot Abnm	p	Pct	UCL	1-Sigma	Case Year
89	21	4	10%	19%	30%	17%	1
90	35	13	10%	37%	25%	15%	2
91	16	4	10%	25%	33%	18%	3
92	20	5	10%	25%	30%	17%	4
93	13	2	10%	15%	35%	18%	5
94	16	6	10%	38%	33%	18%	6
95	8	3	10%	38%	42%	21%	7

Zone B36A: Nonsmoker Only Population

Year	Total	Tot Abnm	p	Pct	UCL	1-Sigma	Case Year
89	8	0	10%	0%	42%	21%	1
90	9	4	10%	44%	40%	20%	2
91	3	1	10%	33%	62%	27%	3
92	2	1	10%	50%	74%	31%	4
93	2	0	10%	0%	74%	31%	5
94	0	0	10%	0%	0%	0%	6
95	3	0	10%	0%	62%	27%	7

Zone B770A: Total Population

Year	Total	Abnormal	p	Pct	UCL	1-Sigma	Case Year
89	20	6	10%	30%	30%	17%	1
90	38	17	10%	45%	25%	15%	2
91	34	4	10%	12%	25%	15%	3
92	37	9	10%	24%	25%	15%	4
93	35	6	10%	17%	25%	15%	5
94	33	7	10%	21%	26%	15%	6
95	21	4	10%	19%	30%	17%	7

Zone B770A: Nonsmoker Only Population

Year	Total	Abnormal	p	Pct	UCL	1-Sigma	Case Year
89	4	0	10%	0%	55%	25%	1
90	12	6	10%	50%	36%	19%	2
91	9	0	10%	0%	40%	20%	3
92	9	1	10%	11%	40%	20%	4
93	8	0	10%	0%	42%	21%	5
94	7	1	10%	14%	44%	21%	6
95	5	0	10%	0%	50%	23%	7

Zone C22I: Total Population

Year	Total	Abnormal	p	Pct	UCL	1-Sigma	Case Year
90	10	3	10%	30%	38%	19%	1
91	12	3	10%	25%	36%	19%	2
92	11	2	10%	18%	37%	19%	3
93	14	6	10%	43%	34%	18%	4
94	7	2	10%	29%	44%	21%	5

Zone C22I: Nonsmoker Only Population

Year	Total	Abnormal	p	Pct	UCL	1-Sigma	Case Year
90	5	2	10%	40%	50%	23%	1
91	7	2	10%	29%	44%	21%	2
92	5	0	10%	0%	50%	23%	3
93	6	2	10%	33%	47%	22%	4
94	2	0	10%	0%	74%	31%	5

Zone C163A: Total Population

Year	Total	Tot Abnm	p	Pct	UCL	1-Sigma	Case Year
89	13	2	10%	15%	35%	18%	1
90	38	12	10%	32%	25%	15%	2
91	50	8	10%	16%	23%	14%	3
92	57	4	10%	7%	22%	14%	4
93	73	8	10%	11%	21%	14%	5
94	90	10	10%	11%	19%	13%	6
95	65	14	10%	22%	21%	14%	7

Zone C163A: Nonsmoker Only Population

Year	NonSmk	Abn NS	p	Pct	UCL	1-Sigma	Case Year
89	8	1	10%	13%	42%	21%	1
90	16	5	10%	31%	33%	18%	2
91	24	4	10%	17%	28%	16%	3
92	24	0	10%	0%	28%	16%	4
93	33	1	10%	3%	26%	15%	5
94	43	2	10%	5%	24%	15%	6
95	28	4	10%	14%	27%	16%	7

Appendix VII: A Proportions Test

The following test is based on the methodology in Hoel's Introduction to Mathematical Statistics (11:156-157):

$$p_S = \frac{514}{2454} \qquad p_N = \frac{287}{2090}$$

$$p_S = 0.209 \qquad p_N = 0.137 \qquad p_S - p_N = .072$$

$$\hat{p} = \frac{801}{4544} \qquad \hat{p} = 0.1763$$

$$\sigma_{pS-pN} = \sqrt{\hat{p}(1-\hat{p})(\frac{1}{2454} + \frac{1}{2090})}$$

$$\sigma_{pS-pN} = 0.011343$$

$$Z^* = \frac{p_S - p_N}{\sigma_{pS-pN}}$$

$$Z^* = \frac{0.072}{0.011343}$$

$$Z^* = 6.34$$

$Z_{0.998} = 3.09$ (11:418) and since $Z^* > Z_{0.998}$ we reject the Null Hypothesis that $p_S = p_N$.

Bibliography

1. American Thoracic Society. "Lung Function Testing: Selection of Reference Values and Interpretative Strategies". Am Rev Respir Dis. March 1991.

2. American Thoracic Society. "Standardization of Spirometry - 1987 Update". Am Rev Respir Dis. March 1987.

3. Arrighi, H. Michael and Hertz-Picciotto, Irva. "Controlling for Time-since-Hire in Occupational Studies Using Internal Comparisons and Cumulative Exposure." Epidemiology. Vol 6, No. 4. July 1995.

4. Bascom, Rebecca, and Ford, Earl . "Don't Just "Do Spirometry"- Closing the Loop in Workplace Spirometry Programs." Occupational Medicine: State of the Art Reviews-Occupational Lung Disease. Vol. 7, No. 2. Hanley & Belfus, Inc. Philadelphia. April-June 1992.

5. Becklake, Margaret R. and White, Neil. "Sources of Variation in Spirometric Measurements". State of the Art Review: Occupational Medicine, Spirometry. pp 2241-260. Volume 8 / Number 2. Hanley & Belfus, Inc. Philadelphia. April-June 1993.

6. Crapo, Robert O. "Pulmonary-Function Testing," The New England Journal of Medicine. Vol 331, No. 1. July 7, 1994.

7. Dockery, Douglas W. "Percentile Curves for Evaluation of Repeated Measures of Lung Function". State of the Art Review: Occupational Medicine, Spirometry. pp 323-338. Volume 8 / Number 2. Hanley & Belfus, Inc. Philadelphia. April-June 1993.

8. Eisen, Ellen A. and others. "Application of Peak Expiratory Flow in Epidemiological Studies of Occupation. State of the Art Review: Occupational Medicine, Spirometry. pp 265-275. Volume 8 / Number 2. Hanley & Belfus, Inc. Philadelphia. April-June 1993.

9. Farzan, Sattar. A Concise Handbook of Respiratory Diseases 3rd ed. Appleton & Lange. Norwalk. 1992.

10. Glindmeyer, Henry W. "Predictable Confusion," Journal of Occupational Medicine. Vol. 23, No. 12. December 1981.

11. Hoel, Paul G. Introduction to Mathematical Statistics. John Wiley & Sons. New York. 1984.

12. Horvath, Edward P. Jr. ed. U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control, National Institute for Occupational Safety and Health. Manual of Spirometry in Occupational Medicine. Cincinnati, OH. Nov. 1981.

13. Knudson, Ronald J. and others. "The Maximal Expiratory Flow-Volume Curve". American Review of Respiratory Disease pp 587-600. Vol 113. 1976.

14. McKay, Roy T. and Horvath, Edward P. Jr. "Pulmonary Function Testing in Industry". Occupational Medicine. Mosby Publishers, 1994.

15. McKay, Roy T. "Spirometry". Occupational Pulmonary Services Center for Occupational Health. University of Cincinnati Medical Center. 1995.

16. Miller, Albert,. "Applications of Pulmonary Function Tests," in Pulmonary Function Tests in Clinical and Occupational Lung Disease. Ed. Albert Miller. Grune & Stratton, Inc. Orlando. 1986.

17. Miller, Albert,. "Prediction Equations and 'Normal'," in Pulmonary Function Tests in Clinical and Occupational Lung Disease. Ed. Albert Miller. Grune & Stratton, Inc. Orlando. 1986

18. Miller, Albert,. "Reference Values for Pulmonary Function Tests," in Pulmonary Function Tests in Clinical and Occupational Lung Disease. Ed. Albert Miller. Grune & Stratton, Inc. Orlando. 1986

19. Monson, Richard R. Occupational Epidemiology. pp. 214-216. CRC Press. Boca Raton. 1990.

20. Montgomery, Douglas C. Introduction to Statistical Quality Control. John Wiley & Sons. New York. 1991.

21. Parker, David L. and others. "A Cross-sectional Study of Pulmonary Function in Autobody Repair Workers," American Journal of Public Health. pp. 768-771. Vol. 81, No. 6. June, 1991.

22. Pennock, Bernard E. and others. "Changes in Measured Spirometric Indices. What is Significant?". Special Communication from the Pulmonary Disease Section, Department of Medicine, University of Oklahoma Health Sciences Center, and Veterans Administration Medical Center. Oklahoma City. July, 1981.

23. Salvan, Alberto and others. "Selecting an Exposure Lag Period." Epidemiology. Vol 6, No. 4. July 1995.

24. Wegman, David H. and Christiani, David C. "Respiratory Disorders," in Occupational Health: Recognizing and Preventing Work-Related Disease. Ed. Barry S. Levy and David H. Wegman. Little, Brown and Company. Boston. 1988.

25. Webster's New Collegiate Dictionary. p 1122. Springfield, MA. G. & C. Merriam Company. 1977.

26. Wright-Patterson LEGACY. pp 2-3. San Diego, CA: MARCOA Publishing, Inc. 1995.

Vita

PII Redacted

Capt Paul W. McAree [redacted]. He graduated from Stevenson High School in 1985 and entered undergraduate studies at Michigan State University in East Lansing, Michigan. He graduated with a Bachelor of Science degree in Mathematics in June 1989. He received his commission upon graduation after completing the two year ROTC program at Michigan State University. His first assignment was at Headquarters, Air Force Military Personnel Center (AFMPC) at Randolph AFB as Officer in Charge of Airman Analysis. He continued on at AFMPC as an Officer Promotion Analyst. While at Randolph AFB, he earned a Master of Science degree in Systems Administration from St Marys University in San Antonio Texas. He completed Squadron Officer School at Maxwell AFB Alabama before entering the School of Engineering at the Air Force Institute of Technology in August 1994.

[redacted]

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	March, 1996	Master's Thesis

4. TITLE AND SUBTITLE
STATISTICAL PROCESS CONTROL IN MEDICAL SURVEILLANCE AN APPLICATION USING SPIROMETRY

5. FUNDING NUMBERS

6. AUTHOR(S)
Paul W. McAree, Captain, USAF

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Air Force Institute of Technology
2750 P Street
WPAFB, OH 45433-6583

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT/GOR/ENS/96M-7

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
74th SGPO (AFMC)
Building 40 5th Street
WPAFB, OH 45433-6583

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This research effort concentrated on applying statistical process control techniques to the results of seven years worth of spirometry exams of Wright-Patterson AFB workers in helping the occupational health squadron identify potential work hazard areas. Each spirometry exam was classified as abnormal or normal based on a comparison with established normals or a significant loss of function from the previous year's exam for the individual. Each test was classified into the work area of the individual and the number of abnormalities per population of the work area was examined. This was accomplished through the use of standardized control charts for fraction abnormal. The base population studied was "out-of-control" for all seven years; however, when excluding smokers only one year of the study for the whole base was "out-of-control". Several work areas were identified as being "out-of-control" and recommended for further study by the occupational health squadron.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Statistical Process Control, Control Charts, Spirometry, Occupational Health, Medical Surveillance	80
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to ***stay within the lines*** to meet ***optical scanning requirements***.

Block 1. Agency Use Only *(Leave blank)*.

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C	- Contract	**PR**	- Project
G	- Grant	**TA**	- Task
PE	- Program Element	**WU**	- Work Unit Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. *(If known)*

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with...; Trans. of...; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

- **DOD** - See DoDD 5230.24, "Distribution Statements on Technical Documents."
- **DOE** - See authorities.
- **NASA** - See Handbook NHB 2200.2.
- **NTIS** - Leave blank.

Block 12b. Distribution Code.

- **DOD** - Leave blank.
- **DOE** - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.
- **NASA** - Leave blank.
- **NTIS** - Leave blank.

Block 13. Abstract. Include a brief *(Maximum 200 words)* factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code *(NTIS only)*.

Blocks 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

★U.S.GPO:1990-0-273-271