

USING TEXTURE TO PREDICT DIAGNOSIS AND DISEASE  
FROM NUCLEAR MEDICINE LUNG PERFUSION SCANS:  
A COMPARISON OF NUCLEAR MEDICINE PHYSICIANS  
TO THE SLOPE OF THE POWER SPECTRUM

by

Mary Virginia Ker

---

Copyright © Mary Virginia Ker 1991

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF PSYCHOLOGY

In Partial Fulfillment of the Requirements  
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

1 9 9 1

## INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

**The quality of this reproduction is dependent upon the quality of the copy submitted.** Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

# U·M·I

University Microfilms International  
A Bell & Howell Information Company  
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA  
313/761-4700 800/521-0600



Order Number 9200015

**Using texture to predict diagnosis and disease from nuclear  
medicine lung perfusion scans: A comparison of nuclear  
medicine physicians to the slope of the power spectrum**

Ker, Mary Virginia, Ph.D.

The University of Arizona, 1991

Copyright ©1991 by Ker, Mary Virginia. All rights reserved.

**U·M·I**  
300 N. Zeeb Rd.  
Ann Arbor, MI 48106



USING TEXTURE TO PREDICT DIAGNOSIS AND DISEASE  
FROM NUCLEAR MEDICINE LUNG PERFUSION SCANS:  
A COMPARISON OF NUCLEAR MEDICINE PHYSICIANS  
TO THE SLOPE OF THE POWER SPECTRUM

by

Mary Virginia Ker

---

Copyright © Mary Virginia Ker 1991

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF PSYCHOLOGY

In Partial Fulfillment of the Requirements  
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

1 9 9 1

THE UNIVERSITY OF ARIZONA  
GRADUATE COLLEGE

2

As members of the Final Examination Committee, we certify that we have read the dissertation prepared by Mary Virginia Ker entitled Using Texture to Predict Diagnosis and Disease from Nuclear Medicine Lung Perfusion Scans: A Comparison of Nuclear Medicine Physicians to the Slope of the Power Spectrum.

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

[Signature]

Date 7/12/91

[Signature]

Date 7/12/91

[Signature]

Date 7/12/91

\_\_\_\_\_

Date \_\_\_\_\_

\_\_\_\_\_

Date \_\_\_\_\_

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copy of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

[Signature]  
Dissertation Director

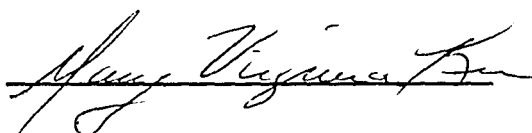
Date 7/12/91

## STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the copyright holder.

SIGNED:

A handwritten signature in cursive script, appearing to read "Mary Virginia", written over a horizontal line.

## ACKNOWLEDGMENTS

I wish to thank my major committee members for their involvement with my dissertation and my graduate career. Dr. William Ittelson, my major advisor, has provided that essential guidance that a graduate student needs in learning to think like a scientist. I greatly appreciate his patience and humor while he tried to keep my feet on the road to finishing my graduate studies, and for introducing me to the joys of physics.

Dr. Terry Daniel provided a strong model of a methodologist, and continually pushed me beyond where I might have gone otherwise. I am grateful to Dr. Daniel for helping me clarify my thinking and explore paths less taken.

I am grateful to Dr. Lee Sechrest for supporting my exploration of medical issues in research.

The late Dr. George W. Seeley deserves special recognition and thanks for his guidance, encouragement, correction, and the occasional kick in the motivator. He supported my work as a graduate student for several years, and provided opportunities for research, professional presentations, and publications. More than anyone else, Dr. Seeley guided my development as a professional. This dissertation is the culmination of a line of research initiated and supported by Dr. Seeley.

In addition, I wish to thank the members of my minor committee for their contributions to my graduate studies. Drs. Judy Dworkin and Michael Bradley encouraged and guided my interest in environmental issues.

Other individuals deserve credit for their support of my doctoral program and research. Dr. Dennis Patton, Dr. Walter Williams, and Dr. Harrison Barrett provided invaluable help in developing the medical and optics sections of this research. Dr. Bob Fiete was responsible for introducing me to Fourier analysis and endlessly patient in explaining it. I thank them all.

Finally, I wish to thank my husband, Nick C. Buchholz, who supported and encouraged me throughout this process, and never lost faith that one day I would finish my graduate studies.

This research was supported in part by grant NCI P01 CA23417 from the National Institute of Health.

## TABLE OF CONTENTS

LIST OF ILLUSTRATIONS.....	7
LIST OF TABLES.....	8
ABSTRACT.....	9
1. INTRODUCTION.....	11
Fractal Geometry.....	11
Fractals and Texture.....	12
Purpose of the Study.....	14
2. FRACTALS.....	17
Fourier Analysis.....	23
Spatial Frequencies and Texture Perception.....	29
Texture.....	29
3. THE LUNG.....	32
The Airways.....	32
The Blood Supply.....	34
The Pleura.....	37
The Structural Support.....	37
Effect of Lung Structure on Texture.....	38
Pulmonary Embolism.....	39
COPD.....	41
CHF.....	42
4. THE NUCLEAR MEDICINE LUNG PERFUSION SCAN.....	46
Creating the Image.....	46
Interpreting the Image.....	49

## TABLE OF CONTENTS (CONTINUED)

5. DECISION MAKING .....	53
Linear Regression.....	57
Signal Detection Theory.....	60
Decision Model Used.....	61
6. METHODS.....	64
Cases.....	64
Images.....	65
Image Analysis.....	67
Subjects.....	69
Procedures.....	69
7. RESULTS.....	71
Interrater reliability.....	71
Relationships between variables.....	73
Relationship of Texture Scales to the SPS.....	74
Relationship of the DCRs to the SPS and Texture..	77
Factor Analysis.....	81
Prediction.....	84
8. DISCUSSION.....	87
Discussion of Results and Conclusions.....	87
Recommendations.....	90
9. APPENDIX A: Beta Weights.....	93
10. APPENDIX B: Abbreviations.....	95
11. REFERENCES.....	96

## LIST OF ILLUSTRATIONS

Table	Page
2.1 Sierpinski gasket . . . . .	17
2.2 Koch snowflake . . . . .	18
2.3 First 3 generations of a Koch snowflake . . . . .	20
2.5 Fractal tree . . . . .	21
2.6 Sine waves . . . . .	24
2.7 Power spectrum of a lung scan . . . . .	27
3.1 Bronchial generations . . . . .	32
4.1 Configuration of gamma camera . . . . .	46
4.2 Views taken by gamma camera . . . . .	50
4.3 Six views of the lungs . . . . .	52
5.1 Lens model of diagnosis . . . . .	56

## LIST OF TABLES

Table		Page
6.1	Cases used in the study . . . . .	64
6.2	The Data Collection Instrument . . . . .	71
7.1	Interrater AUCs . . . . .	73
7.2	Interrater correlations . . . . .	74
7.3	Correlation matrix for all variables . . . . .	76
7.4	Correlations of texture with the SPS . . . . .	78
7.5a	Correlation of DCRs with texture and the SPS .	81
7.5b	Correlation of DCRs with texture and the SPS .	82
7.6	Factor analysis of texture . . . . .	84
7.7	Factor analysis of texture and the SPS . . . . .	85
7.8	AUCs for discrimination of disease . . . . .	86

## ABSTRACT

The lung has been satisfactorily modelled as a fractal, and change in lung structure due to disease is assumed to change the fractal dimensionality of the lung. It is hypothesized that those changes in fractal dimension affect perceptually relevant elements (perceived texture) of the lung, and therefore the fractal dimension may prove to be a predictor of diagnosis. If the fractal dimensionality reflects structure in ways more accurately reflecting changes in lung structure than can be achieved by nuclear medicine physicians, then it may also prove useful as a diagnostic tool.

Fractal dimension is linearly related to the slope of the power spectrum (SPS) as plotted on log-log paper, and the SPS was used as the metric reflecting the fractal dimension. Seventy-two cases were selected that were either normal, had congestive heart failure (CHF), chronic obstructive pulmonary disease (COPD), or pulmonary embolism (PE). Five of the cases had both CHF and COPD. The lung scans from these cases were digitized, with appropriate corrections for linearization, edge artifacts, target nonuniformities and film gamma. Fast Fourier Transforms provided the power spectrum from which the SPS was calculated.

## ABSTRACT (CONTINUED)

Four nuclear medicine physicians read the original lung scans and rated their certainty about the presence of two texture elements, the extensiveness of disease involvement, and presence of the three diseases used (CHF, COPD, and PE).

The results found the SPS to be significantly related to both texture ratings and diagnostic certainty, but inferior as a predictor of disease to either texture rating or diagnostic certainty.

This study reveals the SPS to be a promising but incomplete candidate for machine-algorithm generated diagnosis.

## CHAPTER 1

## Introduction

Fractal Geometry

When Benoit Mandelbrot introduced fractal geometry, he enabled scientists to elegantly describe the structure the physical world often takes. Previous geometries required excruciatingly complicated descriptions to capture the shape of a tree, the form of a cloud, the structure of organic systems. As Mandelbrot pointed out, "Clouds are not spheres, mountains are not cones, coastlines are not circles, and bark is not smooth, nor does lightning travel in a straight line." (Mandelbrot, 1977). By escaping the strictures of integer dimensionality, fractal geometry provides a method of describing these features of nature in parsimonious terms.

Since Mandelbrot's introduction of fractal geometry, numerous researchers have confirmed instances of fractal geometry in the physical world. Coastlines, trees, clouds, and frost crystals all demonstrate fractal geometry (Mandelbrot, 1982). The biological world is equally rich. Liver structure (Cargill, 1989), lung bronchial structure (Goldberger and West, 1987), the arterial system, neurons and the small intestine (Goldberger, Rigney and West, 1990) all appear to be modelled by fractal geometry.

Fractal geometry is more than an interesting quirk of nature. It hints at underlying organizational factors that have shaped almost every aspect of the physical and organic world. Understanding how fractal objects are constructed, and recognizing the significance of changes in their fractal dimension, may allow humans to have a measure of understanding and control of their environment and their own biology never before possible.

#### Fractals and Texture

Given that much of the world appears to be organized according to fractal geometry, and given that humans evolved a perceptual system designed to handle the organization of this world, humans might well perceive the world in ways that are sensitive to changes in fractal geometry.

Psychologists have been slow to investigate this relationship. PsychLit, a computerized version of Psychological Abstracts, only produced 9 citations when asked to search all areas for the word "fractal" for the years 1983-1991. By contrast, Medline, a computerized data base of articles in medical journals, produced 24 citations for 1989 alone. Although the perceptual correlates of fractal geometry have not been extensively investigated, there is some evidence that fractal dimension is correlated with such textural elements as roughness and complexity.

The relationship of the perceived "roughness" of an image to fractal dimension was investigated by Pentland (1984), who had 10 naive subjects rate the "roughness" of curves and surfaces with varying fractal dimensionality. The mean of the subjects' estimates of roughness (on a ten point scale) was nearly perfectly correlated ( $r=.98$ ) with the image's fractal dimension. Using stochastic fractals, Pentland (1984) found fractal dimension and roughness to have correlations of .91.

Investigating the relationship of fractal dimension to complexity, Cutting and Garvin (1987) created line drawings that were self-identical over 2 levels of recursion. Eight subjects rated the images' perceived complexity, and the images with two levels of recursion showed a significant correlation between dimension and complexity ( $r=.82$ ).

Since fractal dimension appears to capture perceptually relevant texture in an image, and since biological organs are often found to have a fractal structure, it is reasonable to hypothesize that changes in fractal dimension that accompany changes in a physiological system can be perceived by humans. Since nuclear medicine physicians (NMP's) describe some diagnostically relevant aspects of the images they work from in terms of texture, the fractal dimension of those images may reflect those relevant textures. This study examines that

proposition in reference to diagnoses based on the pulmonary (lung) arterial bed, which has been shown to be satisfactorily modelled as a fractal (Lefevre, 1983).

#### Purpose of the Study

The purpose of this study is to examine how changes in the texture of scintigraphic lung perfusion studies (which creates an image of the pulmonary arterial bed) affect both fractal dimension and perceived texture, and to discover if medical decision-making could benefit from using more statistical measures of texture than the human perceptual system is able to provide. Specifically, this study asks 3 questions.

1. What is the relationship between texture as perceived by nuclear medicine physicians and texture as described by the mean slope of the power spectrum (SPS) plotted on log-log paper (a measure linearly related to fractal dimension)?

2. To what extent is texture relevant to diagnosis? The relationship of both perceived texture and the SPS will be examined.

For perceived texture, it is possible to develop a linear regression model based on a number of potential perceived texture elements in the form of:

$$X = w_1 \text{Patchiness} + w_2 \text{Focal Defects} + w_3 \text{Extensiveness} \quad (1.1)$$

where  $X$  is the NMP's certainty that a specific disease is present in a given case, 'patchiness' and 'focality' represents the physician's rating of the texture elements usually described as the diagnostically relevant texture components of the image, and  $w$  represents the weights found in the regression procedure. 'Extent' represents how extensive is the disease involvement in the lung, a variable which might be expected to modulate both diagnostic certainty and the SPS, and which will be included in analyses with the texture ratings.

The correlation of texture elements (singly and in combination) to each type of diagnosis will provide a measure of the relationship of perceived texture to diagnosis. The amount of variance in the diagnostic certainty explained by the full linear combination of texture ratings will provide an estimate of how important texture is to diagnosis.

The correlation of the mean SPS (alone and in combination with the standard deviation of the SPS) with the diagnostic certainty ratings will provide a measure of the relationship of actual texture to diagnosis.

3. Which measure better captures the texture relevant to discrimination of diseases: perceived texture ratings, or

the SPS? Since, theoretically, the SPS captures all elements of the texture of the image, while perceived texture may capture only restricted aspects of the image, the SPS has the potential to discriminate diseased from normal cases better than the rating methods used by NMPs.

To answer question #3 both texture ratings and the SPS can be entered into linear discriminant function analysis to discover the best method of differentiating diseased and normals. The canonical variables resulting from the discriminant function analysis can be used to find the probability of correctly identifying a case as normal or abnormal using methods from Signal Detection Theory.

This interdisciplinary dissertation incorporates medical science, fractal geometry, Fourier analysis, gamma scintigraphy, the physics of digitization of film images, decision-making theory under two models, and texture theory. Before proceeding to the methods section of this paper, four chapters will be devoted to explaining those aspects of each of these fields that are relevant to interpreting this study.

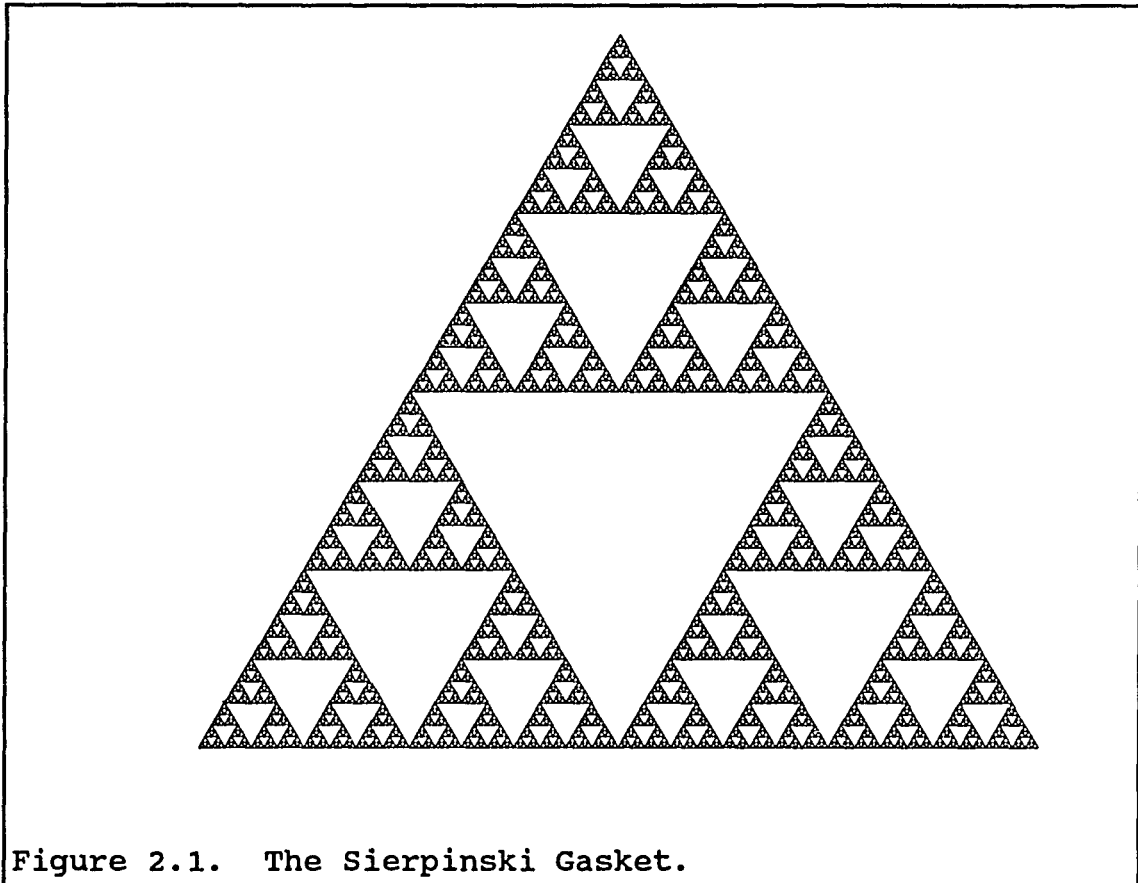
## CHAPTER 2

## Fractals

Mandelbrot (1982) recommends that before trying to understand what a fractal is, one should know what they look like. He then proceeded to present images that looked like natural scenes, but were in fact generated from formulae on a computer. By doing this he demonstrated that the natural world can be convincingly simulated from the simple mathematical instructions used to generate fractals.

Actual objects and scenes in nature, however, may not follow the perfect relations a mathematical formula dictates. This creates a need to distinguish deterministic fractals, defined by mathematical formulae, from stochastic fractals, that appear to follow a fractal geometry, but imperfectly. An example of a deterministic fractal is the Sierpinski gasket (Figure 2.1), which can be reproduced by drawing a point-up equilateral triangle, and drawing lines connecting the middle of each side; then connecting the middle of the sides of each point-up equilateral triangle thus produced, and so on, theoretically, until infinitely small triangles are produced. Stochastic fractals include clouds, trees, mountain ranges and coastlines.

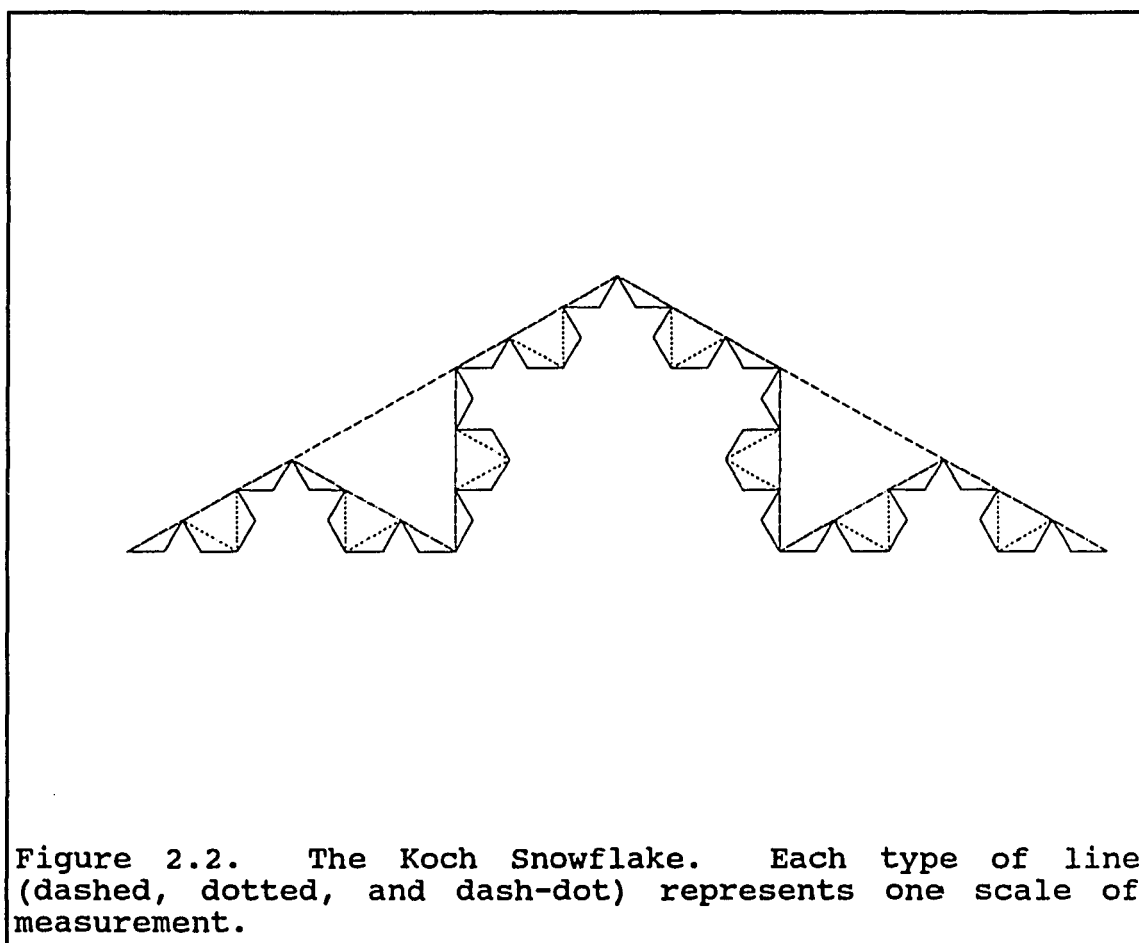
What, exactly, is a fractal? Cargill (1989) explains as follows. In Euclidian geometry, we might measure an object



such as a square. We would use a measuring unit of length  $m$ , and the length  $N$  of the side of the square would be the number of measuring units that fit on that side of the box multiplied by the length of the measuring unit, that is,  $L = N(m) * m$ .

The number of measuring units that fit on the side of the square is inversely proportional to the length of the measuring unit,  $N(m) = m^{-1}$ . This can be demonstrated by halving

the size of the measuring unit, thus requiring twice as many to cover the side. The exponent is determined by the dimension of the object being measured; the side of a square is a line, so its dimension is one.



In fractal geometry we wish to measure irregular objects, such as the irregular curve in Figure 2.2. Using a measuring unit of length  $L$ , the length of the curve is  $2L$ . If we use a shorter measuring unit, one of size  $L/3$ , the

length of the curve can be measured to be  $8L/3$ . If we use yet a shorter measuring unit, the new length will be  $32L/9$ . Thus the measured length of the curve depends on the length of the measuring unit. The smaller the measuring unit, the longer the curve.

The functional relationship between the length of the measuring unit and the length of the line is assumed to be a power law,  $N(L)=L^D$ . For the side of the square,  $D=1$ . For the curve in Figure 2.2,  $m=L$ , and  $N(L)=2$ . For  $m=L/3$ ,  $N(L/3)=8$ , and for  $m=L/9$ ,  $N(L/9)=32$ . These figures can be used to derive a quantitative measure of how a curve scales, thus:

$$N(L)=2=L^D \tag{2.1}$$

$$N(L/3)=8=(L/3)^D \tag{2.2}$$

Dividing Eq. 2.1 by Eq. 2.2 gives

$$\frac{2}{8} = \frac{L^D}{(L/3)^D} \tag{2.3}$$

yielding  $D = 1.2619$ .  $D$  is a quantitative measure of the way the curve scales known as the fractal dimension.

A non-integer dimension between one and two is an object that fills more space than a line, yet less space than a plane. The fractal dimension,  $D$ , is a measure of how well the object fills Euclidian space.

Fractal objects, then, can be seen to be complex objects that will yield different measurements depending upon the scale at which they are measured. However, not every object is fractal that can be so measured. To be a fractal object, the object must also have similar structure (be self-similar) across a range of scales. Deterministic fractals have identical structure (are self-identical) across all scales except one, a condition referred to as being scale-invariant. Thus in a deterministic fractal, the structure at any one scale cannot be distinguished from the structure at any other scale but one.

Self-similarity exists because fractal objects may be defined by an initiator and a generator. The initiator is an initial configuration and a generator is a set of operations that are performed recursively on the initiator. If, for instance, the initiator is an equilateral triangle, and the generator requires that the central third of each side become one side of a projecting

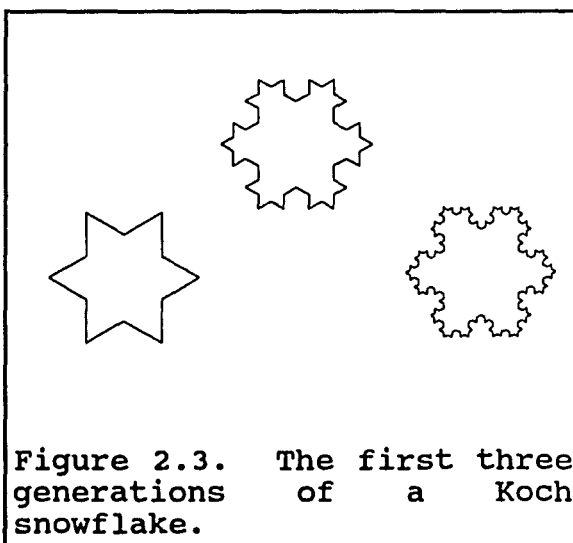
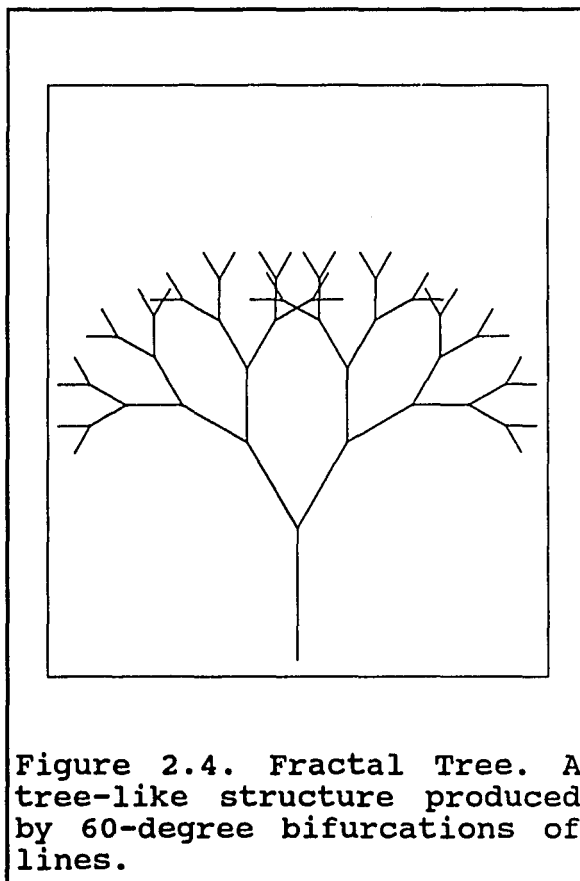


Figure 2.3. The first three generations of a Koch snowflake.

equilateral triangle, one ends up with a Koch snowflake.

Figure 2.3 demonstrates the first three generations of that process. An initiator may be a straight line, and the generator require that one end of that line sprout two lines at a 60 degree angle. Figure 2.4 shows that this quickly leads to a configuration much like a tree.

While deterministic fractals are geometrically identical over all scales except one (that one determined by their maximum scale), stochastic fractals are statistically self-similar over a discrete range of scales. This statistical self-similarity can be quantified in their power spectrum. Power spectra are discussed in the section on Fourier analysis, below. Fractals have a power spectrum that yields a straight line when plotted on log-log paper, and the slope is inversely related to fractal dimension. Cargill (1989) demonstrated the validity of utilizing this method to measure the fractal



structure of nuclear medicine liver scans, that are optically similar to nuclear medicine lung scans, and it is this method of measuring fractal dimension that will be utilized in this study.

#### Fourier Analysis.

In 1822 Baron Jean Fourier demonstrated that a periodic waveform could be analyzed into or synthesized from a linear sum of harmonically related sine waves of specified frequencies, amplitudes, and phases. Since Fourier's discovery it has been shown that this phenomenon is not restricted to periodic waveforms. If all frequencies, and not just the harmonically related ones, are included any nonperiodic waveform can be synthesized or analyzed. In visual terms, this means any image can be perfectly described in a limited set of terms.

For over a century Fourier analysis of visual images was performed infrequently, due to the computation-intensive nature of the procedure. The advent of computers enabled scientists to perform spatial Fourier analysis, and the development of the Fast Fourier Transform (FFT), used to analyze images that have equal sides that are some power of 2, made this analysis fast and relatively inexpensive in computation time.

Although temporal waveforms also exist, discussion of waveforms in this paper shall be restricted to spatial waveforms.

The basic elements of Fourier analysis are sine waves, shown in Figure 2.5. Sine waves can be described by three parameters: frequency, amplitude, and phase. This study assumes phase is not necessary to describe the texturally relevant elements of an image, so it shall not be described further.

The frequency of a sine wave is the number of oscillations the wave performs per unit of distance. If a wave completes two cycles in one centimeter of space, it has a frequency of 2 cycles per centimeter (c/cm). In perceptual research the more usual method of describing frequency is in the number of cycles present per degree of visual angle (c/deg). The more oscillations per unit of distance a sine wave performs, the higher its frequency.

The amplitude of a sine wave is half the distance from its peak to its trough (Figure 2.5). A measure which should be distinguished from amplitude is contrast. Like amplitude, contrast is a measure of the height of the wave form, but differs in that contrast is a relative measure, while amplitude is absolute. The relative nature of contrast is readily apparent in the Michelson contrast, that defines the

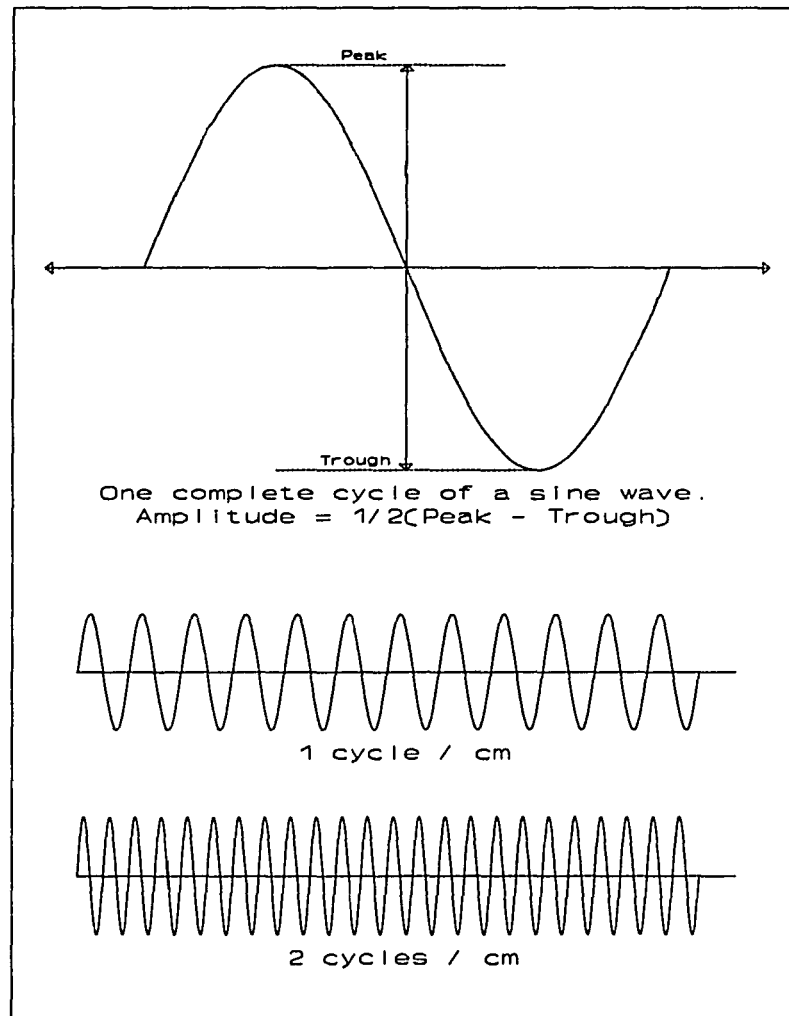


Figure 2.5. Sine waves. Top image is one complete cycle. Middle is series of sine waves at a frequency of 1 cycle per centimeter (not to scale). At bottom is a series of sine waves at twice the frequency of the series at the middle.

contrast of a visual pattern as

$$\frac{L_{MAX} - L_{MIN}}{L_{MAX} + L_{MIN}} \quad (2.4)$$

where  $L_{MAX}$  and  $L_{MIN}$  are the luminances of the peak and trough respectively. Contrast is explained here in order to avoid any possible confusion between the terms, but will not be used in this study.

Waves are sometimes described not by their amplitude but by their power. Power is simply the square of the amplitude. The power spectrum is a histogram giving the power at each frequency. Since the power spectrum will be the basis of the metric used to discriminate diseased cases in this study, its characteristics deserve some discussion.

A cycle cannot be discriminated in less than two pixels. Therefore the lowest frequency discriminable in an image 128 pixels across is  $1/64$  pixels. The overall shape of objects in a lung scan are determined by the low frequency components. They create the large light:dark contrast areas, and therefore can be expected to have the greatest amplitudes, as indeed they do. Texture is not usually considered to be a characteristic of the overall shape of the

object, and so the lowest frequencies are not used in this study.

Higher frequencies contribute detail to the overall shape provided by the low frequencies. In effect, they "fine-tune" the low frequency sine waves, and fewer of them are necessary to create a detailed image, resulting in the negative slope of the power spectrum visible in Figure 2.6. The very high frequencies represent noise, and will not be used in this study. Scintigraphy has an inherent noise level that may vary somewhat depending on the clinical case being imaged due to body size, disease state, and other variables. This noise level is apparent in a power spectrum as a flattening of the power spectrum (Figure 2.6).

Having disposed of the frequencies that provide overall shape, and those that speckle the image with noise (Gaussian noise in the case of a lung scan), what is left are frequencies that represent a range of detailing. If relatively more low frequencies are present, the image will appear smoother, less detailed, and its power spectrum should have a higher slope. If relatively more high frequencies are present, the image will appear to have more detail, and appear rougher, and its power spectrum should have a lower slope.

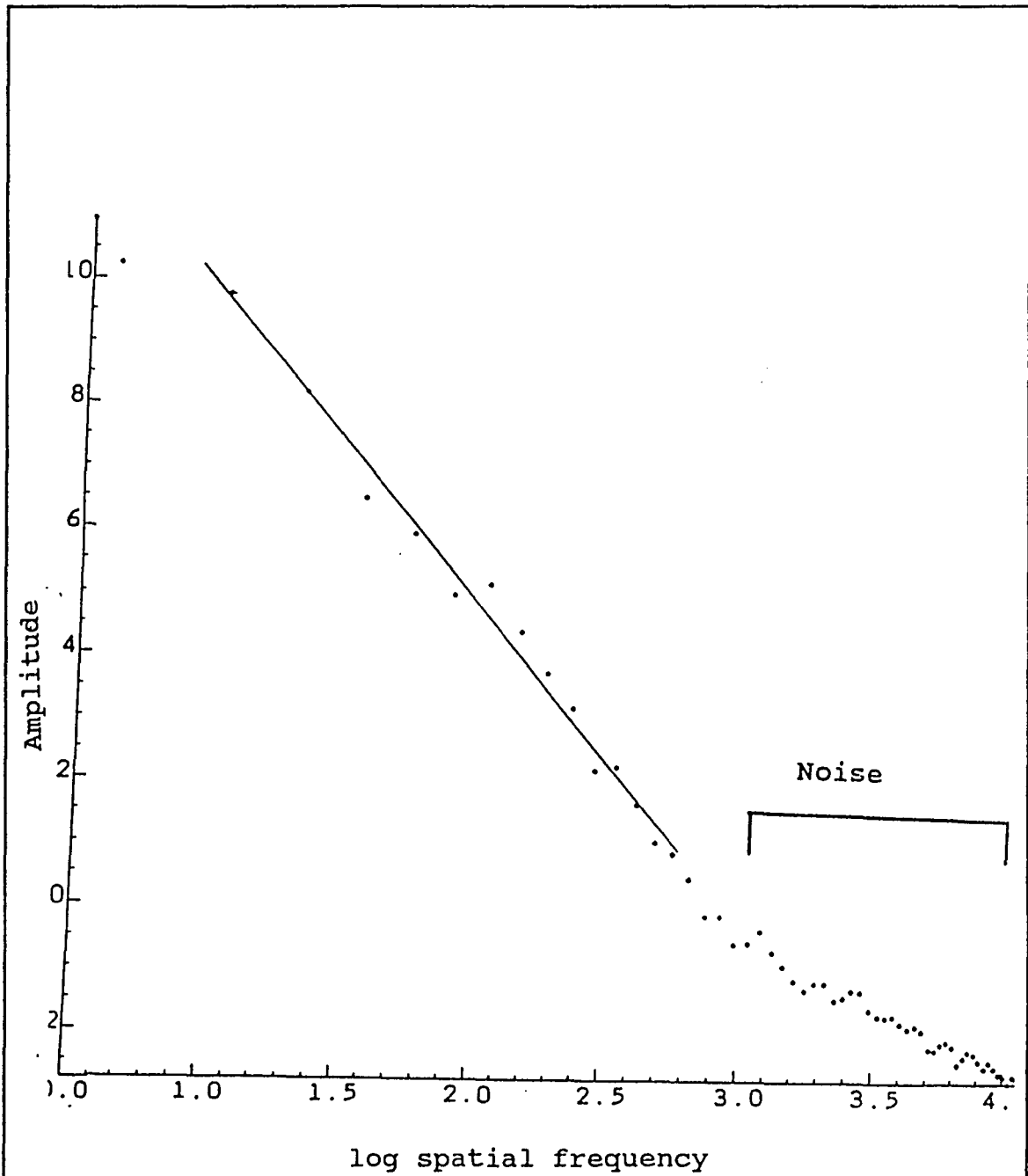


Figure 2.6. A sample power spectrum of a digitized lung scan. The line of best fit through the frequencies used is illustrated, as is the flattened area of noise.

It can be seen from the above discussion that texture can be described in terms of its spatial frequencies, and that the slope of the power spectrum is an appropriate measure for doing so. However, the relevance of spatial frequencies to human texture perception has not yet been discussed.

#### Spatial Frequencies and Texture Perception

Since Campbell and Robson (1969) proposed that the human visual system might contain mechanisms that function as independent, quasi-linear filters tuned to specific bandwidths of spatial frequencies, a great deal of evidence has accumulated to support that model (De Valois & De Valois, 1988; Sutter and Beck (1989), Sagi (1988), Sagi and Hochstein (1984)). This evidence has come from physiological, psychophysical, and anatomical sources. The evidence is reviewed in depth, with great clarity and elegance, by De Valois & De Valois (1988), and will not be reviewed in this dissertation.

#### Texture.

Given that our visual system is tuned to spatial frequency bandwidths, are these used in the perception of texture? Harvey and Gervais (1978) answer in the affirmative. They had 20 subjects examine 30 textures synthesized from various spatial frequencies. The subjects grouped the

textures into 2-5 groups based on their similarity. Harvey and Gervais found that "there is a systematic relationship between the spatial-frequency content of visual textures as represented by activity evoked in four spatial-frequency channels and the perceived similarity of the visual textures."

De Valois & De Valois (1988) extend this conclusion by specifying that textures are those patterns that are characterized solely in terms of the average power at various frequency bands. They assert that for many visual purposes the average spatial frequency (without phase) is the critical component of a scene used in texture recognition. Consider, they ask, a beach; each successive wave redistributes the sand, thus changing the precise spatial waveform of the sand particles, but not its average spatial frequencies. The textural appearance of the sand remains unchanged.

They present evidence to support the model that spatial frequency filtering occurs on a patch-wise, localized basis rather than a global basis, up to the level of the striate cortex. Such a localized analysis would provide information about relative textures at different locations, allowing observers to use it to judge distance and orientation of a broad textural visual field such as a field of grass. A

single global measure of the texture would lose valuable information.

This localized spatial frequency filtering provides the human observer with an advantage over computer-generated Fourier analysis measures such as the slope of the power spectrum. Whether that advantage outweighs the efficiency of the SPS's measure of spatial frequency, that can measure each frequency without degrading it into bandwidths, remains to be seen.

## CHAPTER 3

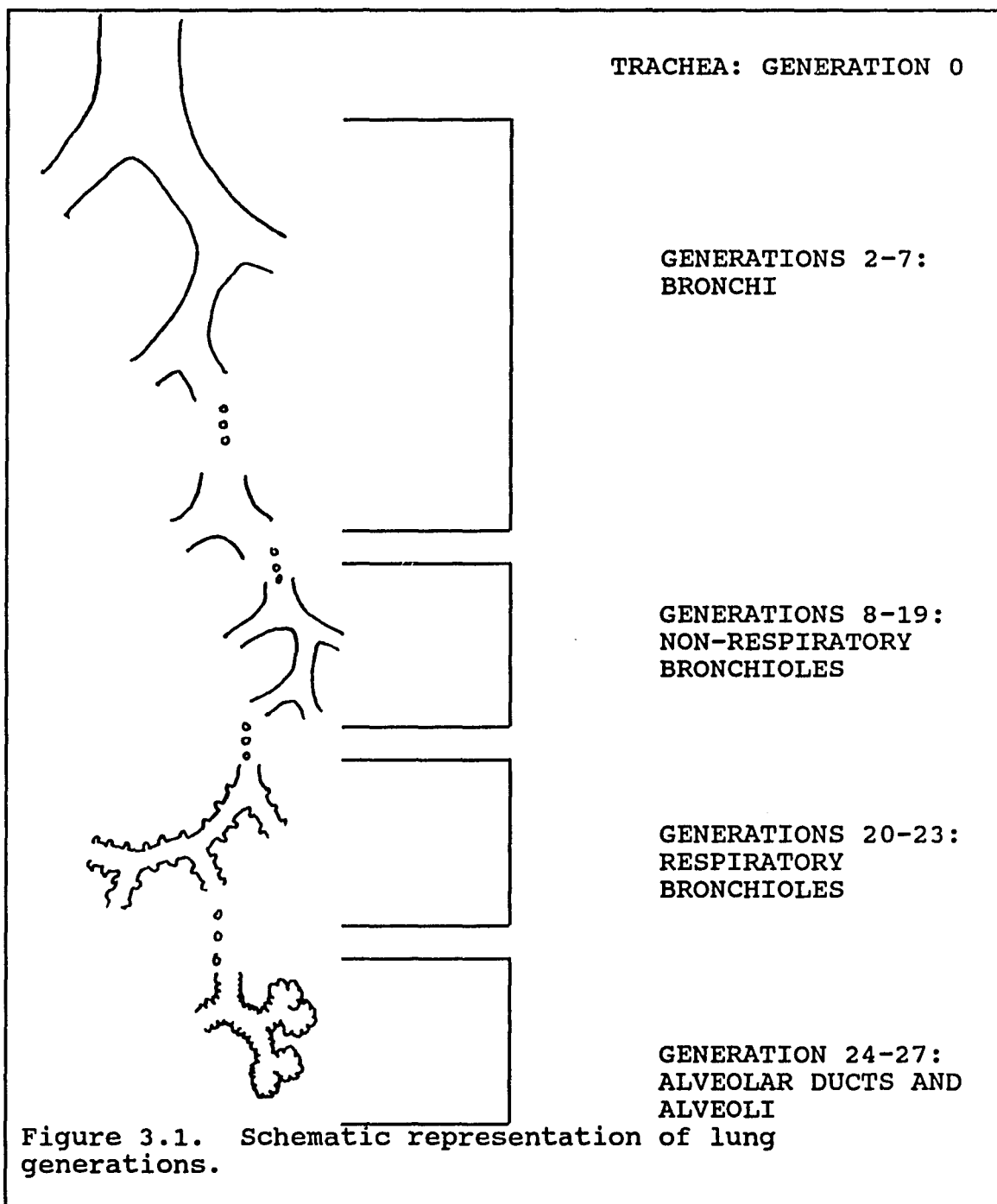
### The Lung

Four components of the anatomy of the lungs are relevant to this study. These are the tracheobronchial airways, the blood vessels, the pleura and the structural support within the lungs.

#### The Airways.

The air conducting system in the lungs, known as the tracheobronchial system, consists of a serially bifurcating tube. Each bifurcation is known as a generation. See Figure 3.1 for a schematic representation of the generations.

The trachea represents generation number zero. For the first through the seventh or eighth generations, the tube is ringed with cartilage and is known as the bronchus. The bronchus serves to conduct gases further into the lungs, and performs non-respiratory functions such as clearing particles out of the lungs. For the eighth through nineteenth generations, the airway loses its cartilage and becomes smaller, membranous tubes called non-respiratory bronchioles. These bronchioles are also part of the gas conduction system. The twentieth to twenty-third generations are called the respiratory bronchioles. These differ from the non-respiratory bronchioles in that they have occasional alveoli along their walls, and thus serve as sites of gas



exchange. Alveolar ducts make up the 24th through 27th generations. Gas exchange actually takes place in the alveolar ducts and alveoli. The respiratory bronchioles, alveolar ducts and alveoli are known as the terminal respiratory unit.

An alveoli may be as little as ten generations from the trachea, or as many as 25, depending on where in the lung it is.

The terminal respiratory unit typically contains 2-5 generations of the respiratory bronchioles, followed by 2-5 generations of alveolar ducts. These ducts branch rapidly, and each duct has openings for 10-16 alveoli. The last of the alveolar ducts opens into 1-3 dome shaped alveolar sacs from which the terminal alveoli project. Alveolar ducts can be viewed as a supporting framework of connective tissue fibers and muscle cells interspersed between a continuous succession of alveoli.

A cluster of 3-5 terminal bronchioles and their terminal respiratory units are known as a lobule. Lobules in humans are partially delimited from each other by connective tissue.

#### The Blood Supply.

The pulmonary arterial system has two branches. The "conventional" pulmonary artery follows the tracheobronchial

airways. The main branches are greater than 1000 microns in diameter, shrinking to between 100 and 1000 microns as it accompanies the bronchioles. As they become about 100 microns in diameter, they are called pulmonary arterioles, and precapillary vessels are only about 35 microns in diameter (Kim & Haynie, 1987). The final segments, called pulmonary capillaries, are 7-10 microns in diameter, and supply the alveolar ducts and alveoli. There are about 280 billion capillaries (DeNardo & DeNardo, 1984).

Because the circulatory and bronchial systems are closely intertwined, conditions that affect an area of the surrounding parenchyma (functioning cellular tissue), such as fluid collection, will affect both systems.

The "supernumerary" arteries do not accompany airways and are smaller than the conventional arteries. They outnumber the conventional arteries and provide about 25% of the total cross-sectional area of the arterial bed near the hilus of the lungs, and about 40% of the total near the periphery. These provide an auxiliary arterial supply to the capillary beds of the terminal respiratory units. These collateral vessels mean that blood supply may be available to a gas exchange unit even if one or more vessels becomes obstructed.

The pulmonary veins drain blood from the capillary bed around the alveoli into the left ventricle of the heart. The pulmonary venous system, along with the left atrium, serves as a blood reservoir for the left ventricle. Along with the pulmonary arterial system the lung has a bronchial arterial system. This latter system receives only a fraction of the cardiac output, and appears to supply lung structures such as the bronchial tree, nerves and ganglia, lymph structures and connective tissue. It also supplies the pleura, the tissue sac around the lungs that separates the lungs from the chest wall.

The bronchial and pulmonary arterial systems have interconnections (anastomoses) in the terminal respiratory units. These connections are few in the normal adult lung, but appear to increase drastically with some lung disorders. Their function is uncertain.

The capillary bed occupies 85-95% of the total alveolar surface. There are approximately 1000 capillaries per alveolus. Therefore the blood flow essentially outlines the alveolar volume. The albumin used in perfusion studies blocks only about 1 in 1000 capillaries, thus on the average, each alveolus is represented in a lung scan by one quantum-emitter.

Gravity and the pressure of the tissue in the lungs creates a pressure gradient within the lungs. This results in greater blood flow near the bottom of the lungs than at the top, a differential of interest in diagnosing congestive heart failure, as will be discussed below.

#### The Pleura.

The pleura is a delicate membrane that surrounds the lung. Each lung has its separate pleura. The pulmonary pleura is that part of the pleura that covers the surface of the lung and dips into the fissures between the lobes. The parietal pleura is that part of the pleura that lines the chest wall. The parietal and pulmonary pleura form a continuous, closed sac. Normally contiguous, the parietal and pulmonary pleura can be pushed apart by collections of fluid, and the area between the two layers is called the pleural cavity.

#### The Structural Support.

The thin tissue of the alveoli is supported in space by a system of connective tissue fibers and a film of surfactant that exerts surface forces at the air-liquid interface of the alveoli. The fibrous matter is made up of elastin and collagen. Each lobe is invested with a fibrous envelope that radiates inward to divide the lobe further into segments. Within the segments are branching networks that invest the

airways and the pulmonary blood system. Short, extremely fine fibers also are found in the alveolar walls. The entire lung parenchyma is supported by a fibrous network that is firmly anchored both at the pleural surface and the hilus.

Effect of lung structure on texture.

The overall texture of the lungs as represented by their blood flow resembles a froth: millions of tiny bubble-like alveoli, each covered by a film of blood. The larger arterioles are largely unrepresented in the perfusion scan, as the blood vessels are not small enough to trap the albumin particles. Only the tiny capillaries covering the alveoli become quanta emitters. The alveoli are represented, on the average, by a single quanta emitter, so what is captured is the location of each alveoli, and these are organized in clusters: a tiny cluster in the alveolus, those grouped in a lobule, lobules grouped into subsegmental units, that are then grouped into segments, the whole grouped into the lung. This replication of grouping patterns at different scales is one factor that has led researchers to characterize the lungs as fractals.

A normal lung will have a relatively smooth appearance on a perfusion scan, slightly darker in the bases than the apex. Diseased lungs may have discrete, localized defects (focal defects) or widespread nonuniformities (patchiness).

They may show a darkening in the apex. The diseases chosen for this study exhibit these three patterns: pulmonary embolism exhibits only focal defects, chronic obstructive pulmonary disease produces primarily patchiness with some focal defects, and congestive heart failure rarely has any focal defects but may change the overall texture of the lungs. The effect of these diseases on the structure of the lung and the resulting lung scan will now be considered in depth.

#### Pulmonary Embolism.

A pulmonary embolism (PE) is any particle of sufficient size to occlude a pulmonary artery. Untreated, PE may be fatal in 6-33% of cases (Dorfman et al.) However, if properly diagnosed and treated, the mortality rate may be as low as 3% of cases (Kim and Haynie).

The effect of an embolus is to block blood from flowing into smaller branches of the pulmonary arterial system. Since the artery and arterioles flow from the midline of the body outward towards the periphery, any blockage creates an area of absent perfusion that spreads in a wedge-shaped defect out from the embolism, and it is this wedge-shaped defect that is looked for when diagnosing a patient for PE. The size of the wedge-shaped defect depends on the size of the artery occluded.

Cutting off the blood flow to an area does not prevent the patient from inhaling; therefore, where the perfusion is blocked, there is rarely a matching blockage of ventilation. A nuclear medicine ventilation scan, that images radioactivity in a gas inhaled by the patient, is often compared to the perfusion scan to discover if there is such a "ventilation-perfusion mismatch". If a defect appears in the same location on ventilation and perfusion scans, this usually means that the problem is not PE, but some other disease.

DeNardo gives the following patterns as being suggestive of PE in a given scintigraph:

1. segmental distribution of well-defined defects.
2. multiple, crescent-shaped peripheral defects or a large avascular area in association with a normal ventilation scan and a relatively normal chest roentgenogram.
3. a generalized reduction in radioactivity in one lung. This may result from incomplete occlusion of the main pulmonary artery, multiple small peripheral occlusions beyond the resolution of the gamma camera, or occlusion of the main bronchus.
4. "Fissural visualization" or a "shrunken lobe" has been attributed to microemboli in a single lobe. Pleural fluid may also cause this.

Wedge shaped defects can be caused by non-embolic vectors. In some cases these can be distinguished from embolic causes by the presence of a "stripe sign": a stripe of activity along the peripheral area of the lung. Small emboli are difficult, if not impossible, to distinguish in the presence of diffuse airways disease (Kim and Haynie). Myocardial infarction or pneumonia may produce defects that are difficult to distinguish from PE (DeNardo).

The textural element contributed by a PE is a discrete shape, contiguous to the outline of the lungs on one or more views. It is not an overall pattern, and might more reasonably be expected to lower the overall amplitude of the power spectrum rather than change its slope.

#### COPD.

Chronic obstructive pulmonary disease (COPD) is marked by uneven regional perfusion, and the defects are not usually segmental (Spencer).

The most common cause of COPD is emphysema. Emphysema is an irreversible disease with destruction of the alveoli and alveolar ducts at the end of the terminal bronchioles (DeNardo). The normally discrete alveoli rupture and coalesce, possibly damaging the capillary bed. The lung becomes overinflated, either in an area or overall. Advanced

emphysema is readily diagnosable, but early stages may not be easy to diagnose.

In emphysema, unlike PE, the ventilation and perfusion defects generally match, i.e., defects are present on both types of scans. Lungs appear long, and there is increased blood flow to the apex of the lung. In the anterior view, overinflation may cause activity to meet across the midline, and the left lower lung may extend below the inferior margin of the heart. Perfusion defects cross anatomic lines. Distribution of perfusion is patchy, and in advanced disease focal defects may appear.

Emphysema is characterized as either "centrilobar" or "panlobar". The centrilobar type is localized in the respiratory bronchiole and is predominantly visible in the lower lobes. Panlobar emphysema is more pronounced in the upper and lateral lung fields.

#### CHF.

Congestive heart failure is a clinical syndrome characterized by disturbances in cardiac output or increased venous pressure (Dorfman, 1974). If the left side of the heart is in failure, the pulmonary veins are unable to discharge into the heart adequately, and become distended. CHF shows nonsegmental, irregular perfusion defects scattered throughout both lungs (Kim and Haynie, 1987). Activity is

relatively increased in the upper lobes. Multiple small PE defects are difficult to distinguish from the small perfusion defects of CHF. In studies by Mohsenifar et al., normal subjects were found to send 53% of the total perfusion to the right lung, with 47% to the left lung. In subjects with CHF, there was no significant difference between the amount sent to the two lungs.

Moreover, there is a reversal of the normal gravity-induced pressure gradient in the lungs. Tanaka et al. (1989) describe the following mechanism to account for the relative reversal of perfusion between the upper and lower zones of the lung in the presence of CHF.

"In the upper lung zone in the normal sitting position, the pulmonary vascular bed is compressed by intra-alveolar pressure to form a non-perfused pulmonary vascular bed where pulmonary perfusion is reduced. When pulmonary venous pressure increased, recruitment of non-perfused pulmonary vascular bed occurs in opposition to intra-alveolar pressure, resulting in increased pulmonary perfusion in the upper zone. As pulmonary venous pressure rises, pulmonary perfusion in the upper zone also increased to make the distribution of pulmonary perfusion uniform, and the nonperfused pulmonary vascular bed disappears. When pulmonary venous pressure continues to further rise, pulmonary perivascular edema begins to develop, starting in the basal lung area. Due to compression, the pulmonary vascular bed decreases in the basal area, followed by a decrease in pulmonary perfusion. As pulmonary perivascular edema continues to develop, the pulmonary vasculature bed in the lower lung

zone also decreases and pulmonary perfusion localizes in the upper zone, with pulmonary perfusion in the lower zone sometimes disappearing."

Pleural effusions may exist. Pleural effusion is the presence of liquid in the pleural cavity (Dorland, 1974). By compressing or displacing the underlying lung, it restricts lung function locally and generally (DeNardo and DeNardo). A large unilateral pleural effusion may cause decreased perfusion in one lung. However, unless major effusions are present, reduced basal perfusion is the only effect on a perfusion scan. (DeNardo & DeNardo in Freeman).

Pleural effusions may lead to the fissure sign: decreased activity along the fissure between the two lobes of the lung (Kim and Haynie) when fluid interposes itself within the major fissures, causing separation of the lobes (DeNardo & DeNardo).

Effusions may result in decreased activity in the left lower lung (DeNardo) and be hard to distinguish from an abnormality of the lungs or an enlarged heart.

If pleural fluid is loculated (restricted in a fixed location), an abnormality will be fixed. But if free-flowing, the abnormality will vary depending on the position of the patient during imaging. In the upright position (as all patient in this study were), the volume of the lung bases

is reduced, and there appears an upward curvature of the lateral lung margin on the anterior and posterior views (DeNardo & DeNardo).

## CHAPTER 4

## The Nuclear Medicine Lung Perfusion Study

A basic understanding of how a lung perfusion study is conducted is necessary to understand the image that is the result. This chapter will provide a simple explanation of the process and instrumentation of creating the lung scan image.

Creating the Image

The nuclear medicine lung scan is created by injecting the patient with radioactively tagged particles (Technetium-99m tagged human albumin) that become trapped in the lungs. These particles remain in the lungs until they are decomposed. As the technetium decomposes (it has half-life of 6 hours (McKeighen, 1980)) it emits gamma quanta. A gamma camera collects the quanta as they leave the lungs.

Figure 4.1 shows a schematic representation of the basic configuration of a gamma camera. Gamma rays may exit the body at any angle. In the most commonly used configuration of collimators, the parallel hole collimator, only those quanta approaching the gamma camera from a near-perpendicular angle are recorded. These are selected by means of a collimator, essentially a thick (about 8 cm) lead plate with closely packed, parallel holes bored through the plate. Each bore is several millimeters in diameter. Gamma quanta approaching at an oblique angle are absorbed by the lead

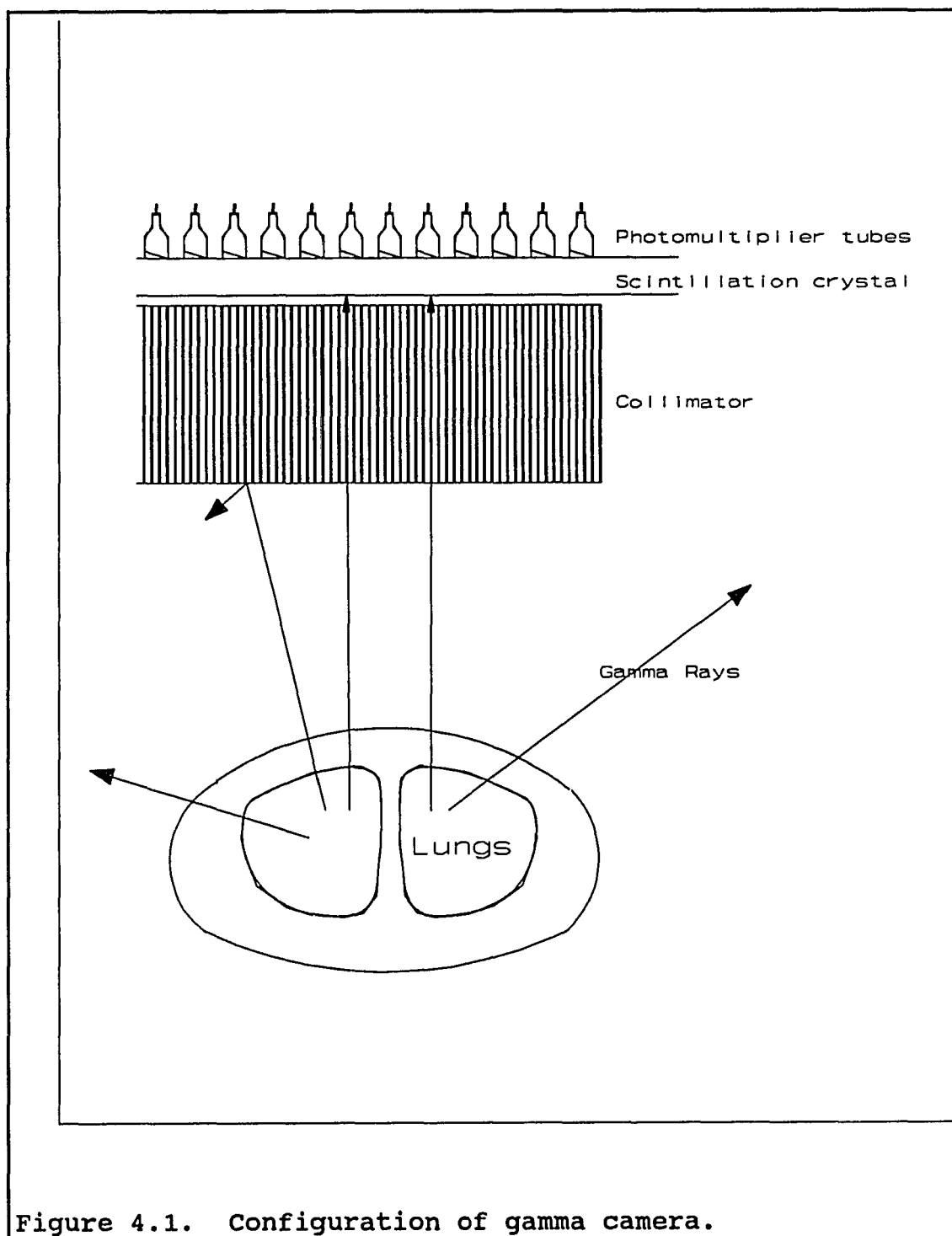


Figure 4.1. Configuration of gamma camera.

shielding each collimator bore.

Those quanta that pass through the collimator strike a scintillation crystal, causing the crystal to scintillate, or emit optical light. The light is recorded by a closely packed array of photomultiplier tubes. Due to the point-spread function of the light, several photomultiplier tubes will be activated by a single quanta, and the most likely x,y location of the quanta must be calculated electronically from the intensity of light recorded by each tube.

Each quanta collected by the camera causes the corresponding x,y location on a piece of film to be briefly exposed. The more quanta collected at that spot, the darker the exposed position on the film becomes. The final product appears as a mass of dots, more or less closely spaced, of varying grey levels.

Quanta can also be collected electronically, with the appropriate x,y location of a quanta incrementing a counter assigned to that area.

The Tc-99m tagged human serum albumin particles are about 10 microns in diameter, and thus cannot pass through the capillaries of the lungs, that are only about 7-10 microns in diameter. (Spencer, 1980). Two to five

microcuries of Tc-99m albumin is usually administered intravenously, containing 60,000 to 150,000 particles (Kim & Haynie, 1987). This obstructs less than 1 in 1000 of the capillaries.

Radioactive particles are lodged wherever there is blood flow (perfusion) in the lung's capillaries. If there is no perfusion, no gamma quanta are emitted from that part of the lung, and the corresponding light (underexposed) area on the film is known as a defect. The relative positioning of light and dark areas on the lung scan are used by the NMP to make deductions about the underlying lung function, and, if possible, its structure.

#### Interpreting the Image

Interpreting this image in three dimensions is a learned skill. The radioactivity is attenuated by the lungs, so more quanta are collected from the side of the lungs near the gamma camera than from the side opposite the camera. This makes the object act as if semi-transparent, and the resulting image is different if collected from the anterior (front) view than if from the posterior (rear) view, but there is some "bleed-through" of quanta from the far side of the lungs. The lung scan is a projection onto two dimensions of a three dimensional object, with all three dimensions represented, in varying degrees, at each point.

Small defects may be difficult to detect, due to the bleedthrough of quanta from normally perfused areas just beyond that spot. However, sometimes one view will allow the NMP to recognize a defect that is not visible on another.

In this study 6 views are used: the anterior, the posterior, the right and left laterals, and the right and left posterior obliques. Obliques are taken at an angle of 45 degrees and 135 degrees to the anterior:posterior axis. (See Figure 4.2) The six views as imaged by the camera are shown in Figure 4.3. The lung is a normal (non-diseased) lung.

Scintigraphic lung perfusion scans are usually obtained to examine the vascular flow to the lungs, and, in conjunction with scintigraphic ventilation studies, provide information about the age of vascular blockages. At present, diagnoses based on these VQ studies rely on qualitative interpretations of the distribution of the tracer, but researchers are looking for quantitative measures which can be calculated from these images to yield more reproducible results.

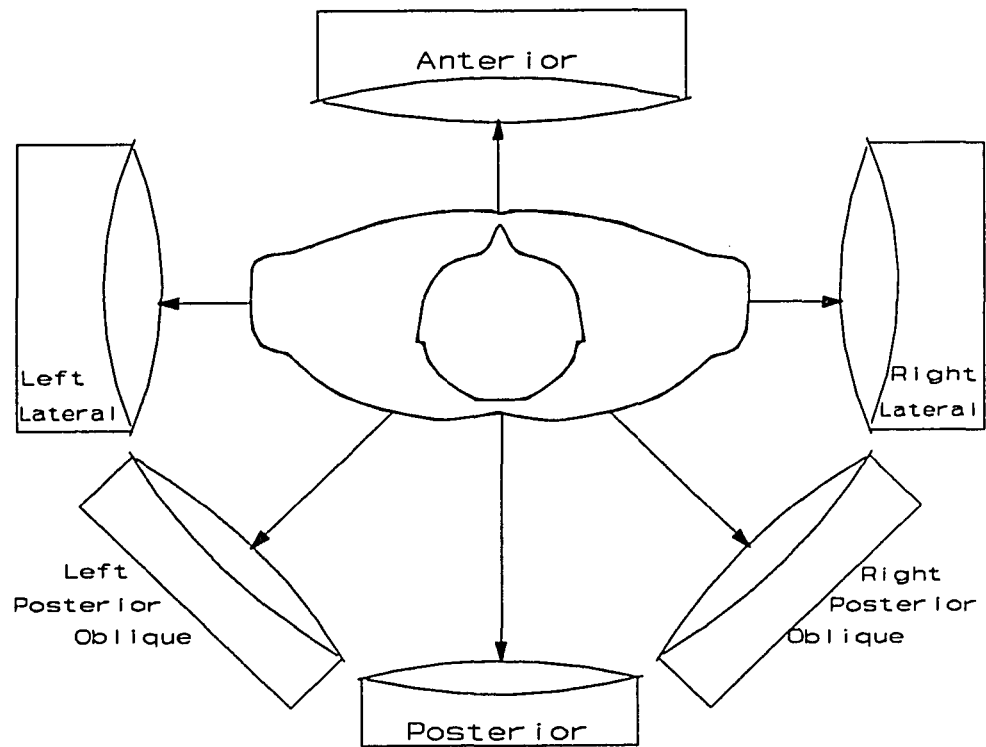


Figure 4.2. Six views taken by the gamma camera.

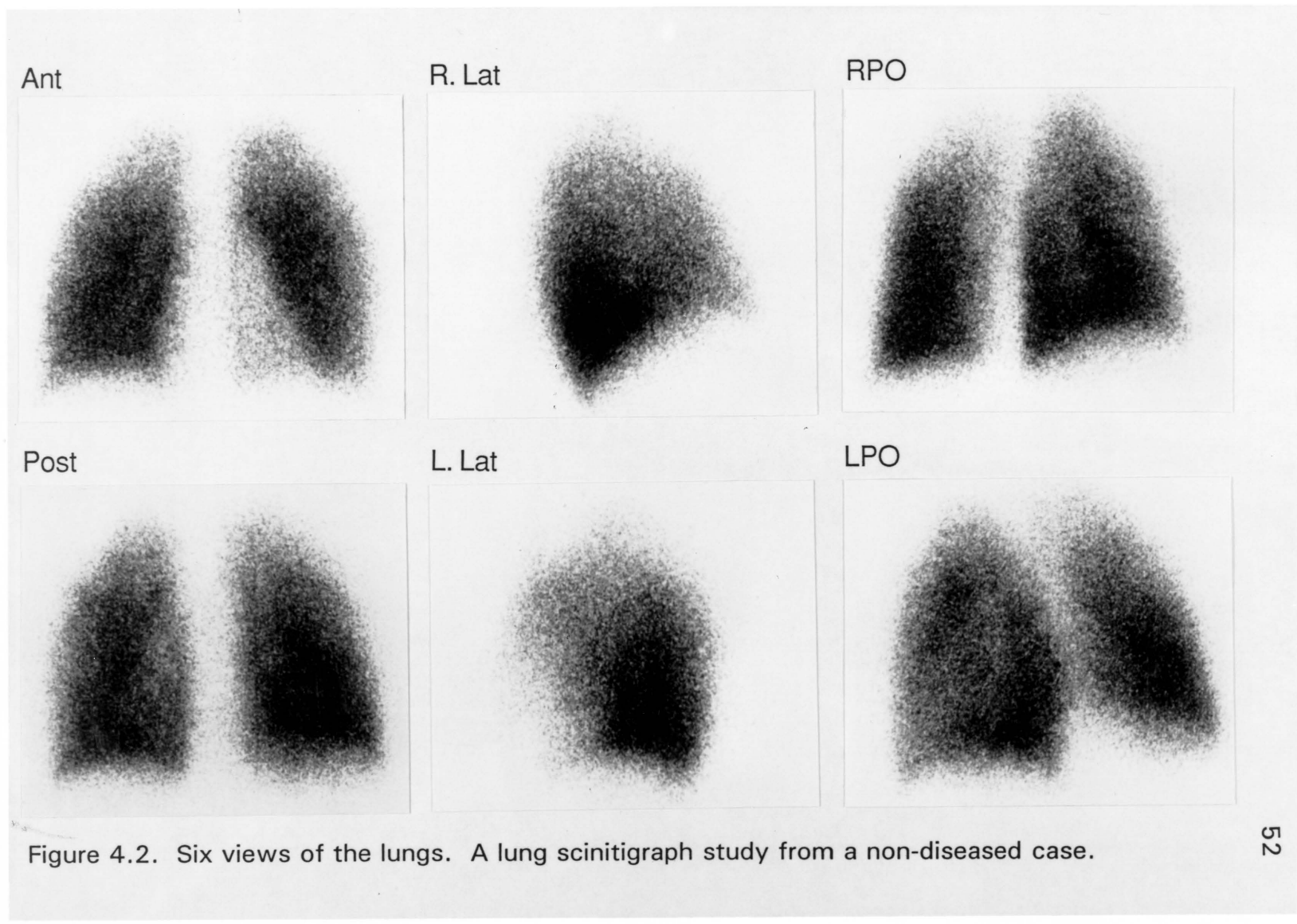


Figure 4.2. Six views of the lungs. A lung scintigraph study from a non-diseased case.

## CHAPTER 5

## Decision Making

Research into decision-making and judgment has proceeded for approximately half a century. Its popularity is partly due to the ubiquity of decision-making in human activities and partly due to concern about the well-recognized lack of reliability in that decision-making.

Blesser and Ozonoff (1972) proposed a model for diagnostic radiologic decision-making with three principal components: psychophysical, psychological, and nosological. Each of these components may give rise to error.

The psychophysical component includes those parts of the process which involve the physics of image creation and image registration (and processing) by the retina and low-level neurons. In order to be diagnostically relevant, physical characteristics of the image must be perceptible by the human. Thus there is little point in increasing resolution of a camera beyond the human eye's resolution, if it is the human eye which is to analyze the image. Error in the diagnostic process may arise from this stage due to limitations in the physical image presented to the radiologist. For example, a lesion in lung tissue in a plain film radiograph is difficult to detect if its location on

film coincides with bone, due to insufficient contrast between the two elements.

The psychological component refers to the creation of meaningful patterns (percepts) from the sensory input. The reduction of sensory input into patterns is an active process which requires the application of context-dependent rules. Such rules lead to an expectational "set" that may induce the radiologist to see what is expected, discarding some features of the image or filling in missing elements.

This phenomenon is demonstrated in eye-tracking studies, where the radiologists fixate a sequence of locations on a film which do not cover the entire radiograph; those parts of the image which are considered significant are viewed carefully and other parts are filled in by the cognitive process.

Errors arise from other aspects of the percept creation process than set phenomena. Radiologic images are inherently ambiguous, and these ambiguities may lead observers to differ in the percept they create of them. For example, one observer may interpret round focal defects in a liver scan as cancer, and another as an artifact arising from gas in the overlying bowel.

Thus interpreting a radiological image involves interaction between visual and nonvisual elements. The

visual elements influence which cognitive rules can be applied, and the cognitive rules influence which visual elements are seen.

The nosological (from "nosology", the taxonomy of disease) element of radiologic diagnosis involves the radiologist's clinical classification of the case. The medical evaluation of a percept is considered to be distinct from the formation of a percept, and errors may occur at either (or both) stages. Such errors may be due to limitations in the radiologist's knowledge or may arise from the nature of classification and diagnosis itself.

Diagnosis arises from both empirical observations and theoretical reasoning, and is the result of inference. For instance, a physician might observe that one lung appears much shorter than the other. If the heart is also enlarged and the gravity gradient reversed, the radiologist might infer that the patient has CHF, and the shortened lung is due to lung displacement from an accompanying pleural effusion. If, on the other hand, other focal defects are present in the lung, a tentative diagnosis of metastatic cancer may be appropriate to explain the lack of perfusion to the base of the lung. A known history of cancer elsewhere in the body increases the probability that cancer is the cause of this feature, but is not direct evidence of cancer.

Thus, even after the radiographic signs are perceived, they must be attributed to the correct underlying pathology, a process which is deductive and inferential and thus subject to error.

The three components of this model are not necessarily sequential; a radiologist may only detect a detail after hypothesizing a specific disease entity and instituting a search for confirmatory signs. The physician may come to the radiograph with a pre-existing expectation of disease based on prior history and examination of the patient which will affect both what areas are looked at and what percepts are formed. These three components should be considered not stages, but relatively concurrent processes for analyzing an image.

All three of the components of this model are considered in the current research. Care was taken to ensure that the optical quality and characteristics of the image analyzed statistically were the same or equivalent to that viewed by the physicians in terms of spatial frequencies and amplitudes (psychophysics). The relationship between textural percepts and physical characteristics was examined (psychology), and diagnostic accuracy measured (nosology). The psychophysics are dealt with in the chapters on fractals, the lung scan,

and methods. In this chapter the components of the creation of percepts and the diagnosis will be discussed.

Those two aspects of this model may be considered jointly as aspects of information integration. Murrell (1977) describes medical decision-making as the process whereby the physician integrates evidence from multiple sources to form a categorical judgment.

There are two broad classes of models to describe such information integration: regression models and Bayesian models (Slovik & Lichtenstein, 1971, in Murrell). Both models are employed in this study; the regression model is linear regression and the Bayesian model is that of Signal Detection Theory (SDT).

#### Linear Regression.

The regression model is derived from Brunswik's (1952) lens model. The lens model, as used in radiologic decision making, can be represented schematically as in Figure 5.1. It proposes that there is some "true" state in the environment, which in radiologic terms means that the patient actually has some disease status (normal or a pathology). The configuration of the disease (its severity, location, distribution, and co-morbidity with other diseases) are also characteristics of the "true" status of the patient.

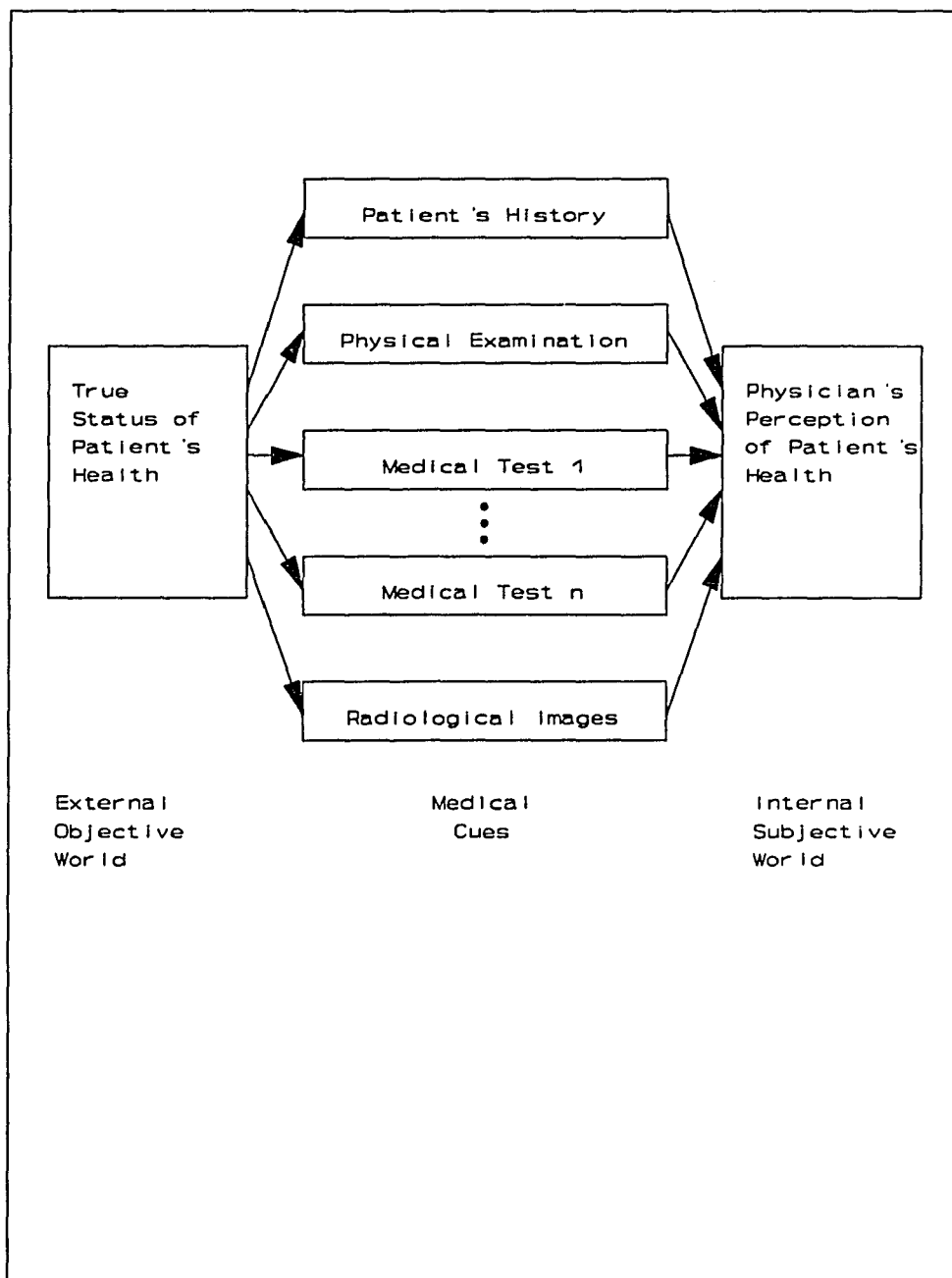


Figure 5.1. Lens model of diagnostic decision making.

The physician, who cannot directly perceive disease, must attempt to decide on whether the patient has a disease through the use of cues generated from the patient: the physical examination, laboratory tests, history, and radiologic images.

The regression model may be used to test a hypothesis that the cues in the lens model can be combined in a weighted linear fashion to explain variance in the final diagnosis. While much of the response variance can often be explained in this fashion, the regression model has been criticized because the weights are arrived at in a post hoc fashion, and are arbitrarily dependent on the particular data entered (Murrell, 1977). Because of this, its generalizability must not be overstated. One attempt to overcome the post-hoc nature of regression is the jackknife procedure (Gray & Schucany, 1972). This procedure is used in the present study in the use of linear discriminant function analysis.

The combination of cues need not be linear, but parsimony demands that when a model of how cues should be combined non-linearly does not exist, a linear model should be examined for its explanatory value.

An application of linear regression is linear discriminant function analysis, which is used to form a weighted combination of values (a canonical value) to

optimally discriminate between two (or more) groups. In this study the canonical variables from linear discrimination will be used to construct a ROC (Receiver Operating Characteristic) curve.

The regression model is similar to that of Blesser and Ozonoff's in its recognition that the physician must integrate information to arrive at a decision. It can describe the integration process, but does not deal with the nosological accuracy of the decision. That is left to Signal Detection Theory.

#### Signal Detection Theory.

SDT was originally developed as a mathematical model to aid in detecting radar signals (Peterson et al., 1954). It has found use since as a model to help understand the behavior of human observers in detecting visual and auditory signals (Tanner & Swets, 1954), in medical diagnostic tasks (Swets, 1969, and others too numerous to mention), and many other fields where accuracy was of interest.

Originally developed to model detection of a simple signal in a noisy background, its premise was that the noise-only and the signal-plus-noise distributions generated internal representations (evidence values) that had different mean values. The observer is assumed to compare the evidence value of the stimulus being evaluated to some internal

criterion in order to assign it to a distribution (Murrell, 1977). If the criteria are represented as a series of certainty thresholds, representing gradated certainty on the part of the observer, an ROC curve can be generated for a set of data. An ROC curve is the plot of true positive decisions versus false positive decisions at each criterion threshold. The area under the ROC curve (AUC) is equal to the probability in a two-alternative forced-choice situation that the observer will correctly detect a signal-present (positive) case.

When ROC analysis was applied to medical decision making by Metz (1978), the simple signal detection task was replaced with a model of diagnostic combination of evidence. As Metz pointed out, diagnostic test data may fall in a continuum of values rather than discrete categories. A series of thresholds can be applied to this data to yield an ROC curve.

ROC analysis need not deal with the internal evidence values that generate the continuum that is subdivided by certainty thresholds. It's primary use in medical decision making research is to evaluate the nosological accuracy of physicians.

#### Decision Model Used.

This study uses the regression model to examine how the percepts of texture and disease extent are related to the

physical image characteristic of texture, and the signal detection model to discover if the SPS is equal to or better than the combined texture judgements of the physician in distinguishing the disease status of a patient.

The actual decision-making situation is being modelled in this study as follows.

The texture of the scintigraph is a clinically relevant stimulus of the distal environment, and this stimulus can be adequately represented by the SPS.

Texture stimuli drive an evaluation process that produces a texture rating, and that rating is an evidence value. These texture evidence values are combined using a linear weighted scheme (linear regression) to predict a diagnostic response value. It is not necessary to capture every evidence value; the correlation coefficient,  $r^2$ , can estimate the proportion of the diagnostic response value variation due to the variation in the texture rating stimulus values.

This provides a model of diagnosis (nosological decisions) in terms of texture ratings (percepts). The beta weights of the regression equation estimate the contribution of each percept to the diagnosis.

The relative ability of the SPS versus the texture ratings as a diagnostic metric can be estimated by a combination of linear discriminant analysis and ROC analysis. Linear discriminant analysis will disclose which set of scores can be used to better discriminate two groups (normals versus diseased); ROC analysis yields the probability of making a correct classification based on the linear discriminant score.

## CHAPTER 6

## Methods

Cases.

Seventy-two diseased and normal cases have been selected on the basis of verification at the levels of Classes I-IV in the Ker et al. [1988] verification protocol. This protocol requires that verification be provided by either autopsy, biopsy, direct inspection during surgery, or agreement among two or more imaging techniques. Table 6.1 describes the breakdown by disease of cases used in this study.

These cases are either normal or have one of three diseases: CHF, COPD or PE. Five cases have CHF co-morbid with COPD. These diseases were chosen for study both because they can be roughly ordered along a systemic:focal dimension, and because of availability. The ordering of those diseases along the systemic: focal dimension was defined by a nuclear medicine physician, with CHF being the most systemic and PE being the most focal.

These cases represent all the cases having these diseases (or being normal) which were (1) given a perfusion scan between December, 1982 and June, 1990 at the University of Arizona Health Sciences Center (2) imaged on the Picker 4/15 gamma camera, (3) had no observable artifacts such as

pacemakers, and (3) were verifiable by Classes I-IV of the Ker et al. (1988) protocol.

DISEASE PRESENT	N
CHF	10
COPD	32
PE	15
Normals	20

Table 6.1: Cases used in study.

#### Images.

All lung studies were done at the Nuclear Medicine Department of the University of Arizona Health Sciences Center. Lungs were imaged using a Picker 4/15 Gamma camera with a fifteen inch field of view and parallel hole collimator. Data acquisition was controlled by a PDP 11 34 computer with Gamma-11 software. Gamma quanta were collected until 50,000 counts were recorded in a test patch of the lungs.

Six views were collected for each lung scan: anterior, posterior, right and left laterals, and right and left posterior obliques.

The scans were collected on film, and digitized using a Plumbicon camera, an 8 bit A/D converter, and a frame grabber to record the digitized images for storage on a VAX 8600. A pixel was represented by one byte, permitting 256 grey

levels. Images have been corrected for linearization of the camera, nonuniformities in the camera's image field, edge artifacts, and the gamma value of the film used to originally record the lung scans. No MTF correction was necessary as all lung scans were performed on the same gamma camera and system.

Each 512x512 pixel image was reduced to 256x256 pixel size to eliminate edge artifacts in the frequency domain. No portion of the lung scan was sacrificed in doing so. The image was then further reduced to 128x128 pixels through subsampling to reduce time needed to calculate power spectra. Subsampling was performed by selecting every other pixel, on every other row. The slopes of the power spectra (SPS's) of the subsampled images were identical to the SPS's of the larger images in a test set of 5 randomly selected images.

Several corrections to the pixel values were necessary. In the system used in this study, the optical density of the object to be digitized had a linear correlation of .976 with the grey level represented by that density in the digitized image, and the digitized grey level values were corrected to closer represent the optical density value. Moreover, the Plumbicon camera, being based on tube technology, has a nonuniform field. Even if it is used to take a picture of a perfectly uniform field, nonuniformities will appear. It was

found that these nonuniformities can shift position and number in a matter of minutes. Therefore an image of a blank field was recorded every few frames, and these images were used to correct the nonuniformities present in the lung scan images. The gamma value of the film used to record the lung scans in the gamma camera system was obtained from its manufacturer, and the pixel values corrected for gamma value.

The gamma camera system used also collects four 15-second periods of quanta emission from the posterior view directly into digital memory. Comparison of the power spectrum of this digital images with the corrected digitized images showed power spectra that were almost identical except for the expected difference in amplitude, confirming the accuracy of the corrections, and the validity of the digitization process.

#### Image Analysis.

The SPS depends on which frequencies are used to calculate it. As described by Cargill (1990), the rationale for using the SPS as a discriminant is that the lungs have an approximately fractal geometry, and as that geometry is affected by disease, the SPS changes. Fractal objects show a straight line on a log-log plot; therefore that part of the power spectrum which demonstrates a straight line is the appropriate area over which to calculate a slope. We found

that frequencies from  $e^{1.1}$  to  $e^{2.7}$  formed a straight line on a log-log plot. Pearson's  $r$  about the line of best fit through the power spectrum was uniformly high; the absolute value of  $r$  had a mean of .98, and standard deviation of .009. Ninety-seven percent of the absolute value of  $r$ 's were over .97.

The average radial power spectrum for each of the six views was found, and the line of best fit calculated from the power spectrum. The slope of this line was found, and the mean slope across the six views used as the measure of the SPS, as well as the standard deviation of those slopes.

Cargill (1989) had some success using the mean and standard deviation of the SPS of six views to discriminate normal and abnormal cases. Her method of classification was arbitrary and was not used here. Instead, the mean and standard deviation were classified with the stepwise linear discriminant function analysis program of BMDP. Ordering cases by their linear discriminant canonical variable value permitted construction of a Receiver Operating Characteristic (ROC) curve. In addition, linear discrimination was used to find canonical variables for the texture ratings and that canonical variable used to calculate an ROC curve.

Predictive ability of the linear discriminant functions was maximized through use of a jackknife procedure in BMDP. In addition, the mean SPS alone was ordered and used to calculate an AUC.

The data underlying the ROC curves were not assumed to be normally distributed, so the nonparametric, trapezoidal-rule method of estimation from Hanley and McNeill (1982) was employed to calculate AUCs and their standard errors. Significance between two independent ROC curves was tested using the method from Massof & Emmel (1986).

#### Subjects.

Four nuclear medicine physicians (NMPs) volunteered their time to evaluate the images. All NMPs were very experienced, with years of experience in nuclear medicine ranging from 6 years to over 20 years.

#### Procedures.

To conduct the NMP evaluation, all images were taken to the NMPs' offices where they were read under each doctor's normal viewing conditions. The images viewed by the nuclear medicine physicians were the original lung scan films, i.e., full size, full resolution images. Names and case numbers were taped to be invisible to the physicians. Two NMPs read them in the presence of the experimenter; two read them alone. Physicians rated each image presented on their

certainty about the presence of CHF, COPD, and PE using the same six-point scale to rate each disease. They also rated the case on two texture variables commonly used by NMPs to describe lung scans, as well as estimating how extensive the disease involvement of the lung is.

The scale used is a commonly used scale developed for ROC analysis and medical psychophysics. Table 6.2 contains the data collection sheet, with all scales used.

0	absolutely certain this condition is absent					
1	relatively certain this condition is absent					
2	just guessing this condition is absent					
3	just guessing this condition is present					
4	relatively certain this condition is present					
5	absolutely certain this condition is present.					
-----						
	0	1	2	3	4	5
Presence of CHF:						
COPD:						
PE:						
any other disease:						
patchiness:						
focal defects:						
How extensive is the disease involvement in this patient's lungs?						
None			Minimal			Moderate
estimated area:	<10%			11-50%		Severe
					51-80%	Total
						>80%

Table 6.2: Instrument used by physicians to rate lung cases.

The disease scales allow the experimenter to calculate 5 points along an ROC curve for each disease and each texture variable.

## CHAPTER 7

## Results

The results from this study are organized around three rubrics: interrater reliability, the relationship between variables, and the ability of variables to predict disease. They will be addressed in that order. Beta weights are reported in Appendix A for all correlations.

Interrater reliability.

Interrater reliability was examined under both models. Under the Bayesian model, it was examined by determining if ROC curves were significantly different between NMPs. Table 7.1 contains the AUCs and their associated standard errors. Given these sample sizes, any statistical test of differences will be low in power (Hanley and McNeill, 1982), but with the exception of NMP4's CHF AUC, no two AUCs were significantly different.

AUCs:	NMP1	NMP2	NMP3	NMP4
COPD:	.81±.06	.92±.04	.87±.05	.88±.06
CHF:	.87±.07	.97±.03	.92±.05	.65±.10
PE:	.99±.01	.97±.03	.96±.03	1.00±.00

Table 7.1. AUCs and their standard errors for each NMP for each disease type. Test sets consist of one disease type plus true normals.

Interrater reliability was also examined under the regression model. Table 7.2 contains the correlation

matrices of NMPs on each rating scale with the z-standardized mean correlation for each scale. The highest mean

CHF	NMP # 1	NMP # 2	NMP # 3	mean
NMP #1	1.0000			.547
NMP #2	.6364	1.0000		
NMP #3	.4875	.6551	1.0000	
NMP #4	.4299	.4766	.5644	
COPD	NMP # 1	NMP # 2	NMP # 3	
NMP #1	1.0000			.602
NMP #2	.5527	1.0000		
NMP #3	.5688	.6865	1.0000	
NMP #4	.6133	.5782	.6060	
PE	NMP # 1	NMP # 2	NMP # 3	
NMP #1	1.0000			.856
NMP #2	.8530	1.0000		
NMP #3	.8264	.8650	1.0000	
NMP #4	.8184	.8944	.8747	
OTHER	NMP # 1	NMP # 2	NMP #3	
NMP #1	1.0000			.420
NMP #2	.2308	1.0000		
NMP #3	.3917	.4306	1.0000	
NMP #4	.3268	.5359	.5635	
PATCHY	NMP # 1	NMP # 2	NMP # 3	
NMP #1	1.0000			.546
NMP #2	.4815	1.0000		
NMP #3	.4731	.7472	1.0000	
NMP #4	.5731	.4791	.4559	
FOCAL	NMP # 1	NMP # 2	NMP # 3	
NMP #1	1.0000			.630
NMP #2	.6157	1.0000		
NMP #3	.7499	.5702	1.0000	
NMP #4	.6852	.4037	.6925	
EXTENT	NMP # 1	NMP # 2	NMP # 3	
NMP #1	1.0000			.747
NMP #2	.6791	1.0000		
NMP #3	.7348	.8041	1.0000	
NMP #4	.7677	.7249	.7551	

Table 7.2. Interrater correlation on each rating scale. Mean of standardized regression coefficients are provided for each rating scale.

correlation between raters was on the PE DCR ( $r=.856$ ), which is not surprising given their very high AUCs for that disease (Table 7.8). The lowest reliability was on the "other" scale ( $r=.42$ ), and this scale will not be used in further analyses. The other scales vary from  $r=.546$  ("Patchy") to  $r=.747$  ("Extent").

Such relatively low reliability between physicians is very common, and is one of the motivating influences behind research into alternative metrics to establish diagnoses (Kundel & Revesz, 1974).

#### Relationships between variables.

The variables used in this study are the two texture rating scales, the extent scale, the three DCRs, and the mean and standard deviation of the 6 SPS scores. Individual relationships were examined through correlation analysis, and overall relationships examined through factor analysis.

#### Correlation Matrix.

Table 7.3 contains the correlation matrix for all cases used in the study. The correlations between disease rating and the SPS are all significant ( $p<.006$  on all comparisons), as are correlations between texture ratings and the SPS ( $p<.001$  on all comparisons).

The relationship between disease rating and texture rating is consistent with NMP perceptions. The presence of

focal defects has its highest correlation with PE ( $r = .683$ ) and patchiness has its highest correlation with COPD ( $r = .607$ ). CHF has no high correlation with any one texture element, but appears almost equally related to patchiness and focality.

	SPS	s	CHF	COPD	PE	PATCHY	FOCAL
SPS	1.000						
s	.055	1.000					
CHF	.176	.213	1.000				
COPD	.411	-.048	.247	1.000			
PE	.162	.086	.118	.067	1.000		
Patchy	.347	-.037	.364	.607	.044	1.000	
Focal	.187	.059	.312	.258	.683	.138	1.000
Extent	.477	.032	.199	.405	.363	.461	.249

Table 7.3. Correlation matrix for the mean slope of power spectrum, the standard deviation of the SPS's (s) and rating scales. All cases are included in the analysis.

#### Relationship of Texture Scales to the SPS.

Multiple regression was performed to examine how the texture rating scales were related to the SPS. Table 7.4 presents the correlation of various linear combinations of texture ratings with the SPS. Case sets consisting of one disease type plus the 20 true normals were examined. In addition, all cases were analyzed as a group to examine the correlation of the SPS with texture across disease types. All correlations were significant.

variables:	R	R <sup>2</sup>	St. Err.	F	p
CHF:					
Patchy	.50514	.25516	.19733	47.28	.0000
Focal	.35332	.12483	.21390	19.68	.0000
Extent	.55301	.30582	.19050	60.80	.0000
Patchy & Focal	.54126	.29296	.19296	28.38	.0000
Patchy & Extent	.57516	.33081	.18772	33.86	.0000
Focal & Extent	.58966	.34770	.18534	36.51	.0000
All 3 Variables	.60165	.36199	.18397	25.72	.0000
PE:					
Patchy	.26812	.07189	.20065	10.92	.0012
Focal	.17476	.03054	.20507	4.44	.0368
Extent	.41968	.17613	.18904	30.14	.0000
Patchy & Focal	.29334	.08605	.19982	6.59	.0018
Patchy & Extent	.42790	.18310	.18891	15.69	.0000
Focal & Extent	.42613	.18159	.18909	15.53	.0000
All 3 Variables	.43347	.18790	.18903	10.72	.0000
COPD					
Patchy	.42120	.17741	.20010	45.29	.0000
Focal	.34684	.12030	.20693	28.72	.0000
Extent	.52198	.27247	.18818	78.65	.0000
Patchy & Focal	.48410	.23435	.19351	31.98	.0000
Patchy & Extent	.56384	.31791	.31139	48.71	.0000
Focal & Extent	.56384	.31791	.18265	48.71	.0000
All 3 Variables	.57392	.32938	.18154	34.05	.0000
ALL CASES					
Patchy	.34741	.12069	.21590	40.22	.0000
Focal	.18668	.03485	.22620	10.58	.0013
Extent	.47676	.22730	.20239	86.19	.0000
Patchy & Focal	.37461	.14033	.21384	23.83	.0000
Patchy & Extent	.49804	.24804	.20000	48.16	.0000
Focal & Extent	.48188	.23221	.20209	44.16	.0000
All 3 variables	.50243	.25243	.19976	32.75	.0000

Table 7.4. Correlation of texture variables with the slope of the power spectrum. The first three groups represent the correlation within a set consisting of the indicated disease type plus normal cases. The last group represents the correlation within a set consisting of all three disease types plus the normal cases.

The smallest relationship between texture ratings and the SPS belonged to cases with focal disease, and the largest relationship to the cases with the most systemic disease. The mediating variable in this relationship may be the extent of the disease. The texture ratings reflect certainty that a texture exists, and certainty about texture is weakly correlated with the extent of a disease ( $r=.25$  to  $.46$ ). An NMP may be as certain focal defects exist when only a single defect is present as when multiple defects exists, but is more likely to be certain that a systemic disease is present if systemic effects are present. Since the SPS analyzes the image as a whole, the more of the lung field the disease covers, the more the SPS should measure the texture being described by the NMP. If the NMP is rating a systemic pattern, this systemic pattern will have a larger impact on the SPS and so the texture rating should have a higher correlation with the SPS. If the NMP is rating a localized pattern, that pattern will have little effect on the SPS and so the texture ratings and the SPS should have a lower correlation. That is what we see. The correlation of patchiness to the SPS ranged from  $.27$  (PE) to  $.51$  (CHF), and the correlation of focal defects to the SPS ranged from  $.17$  (PE) to  $.35$  (CHF).

This explanation is supported by the fact that when only one disease type is considered, the variable with the highest correlation with the SPS is the extensiveness of the disease; the second highest is patchiness, and the third is the presence of focal defects.

This can clearly be seen from the linear models.

$$\begin{aligned}
 SPS_{CHF} &= .17402_{PATCHY} + .18796_{FOCAL} + .37553_{EXTENT} \\
 SPS_{COPD} &= .13462_{PATCHY} + .20429_{FOCAL} + .38664_{EXTENT} \\
 SPS_{PE} &= .08974_{PATCHY} + -.0839_{FOCAL} + .42532_{EXTENT} \\
 SPS_{ALL} &= .16026_{PATCHY} + .06842_{FOCAL} + .38590_{EXTENT}
 \end{aligned}$$

The 2-variable and 3-variable correlations on Table 7.4 show that the texture ratings contribute little beyond that of the extent rating in explaining the SPS. The greatest contribution is in CHF, where 7% more variance in the SPS is explained by adding texture to extent in the equation.

#### Relationship of the DCRs to the SPS and Texture scales.

Tables 7.5a and 7.5b show the correlation of DCRs with the other variables. Tables 7.5a and 7.5b (formatted as two tables for the sake of spacing) contain data from test sets consisting of the true normals plus one disease type.

Correlations were higher between the two types of perceptual judgements (DCRs and texture ratings) than between

the SPS and either the texture judgments (shown in Table 7.4), or the DCRs. Adding the SPS to the linear combination of texture ratings (and extent ratings) did not improve correlations with DCRs.

Sixty-eight per cent of the variability in diagnoses of PE could be explained in terms of the judgments about patchiness, focality and the extent of disease involvement; adding the SPS did not change that figure. Although less of COPD and CHF could be explained by texture ratings, the SPS did not improve those correlations by more than one point, indicating that the information the SPS provides the physician is either already incorporated into the texture ratings or irrelevant to diagnosis.

In fact, most of the variability in diagnostic certainty could be explained by two variables. In cases of CHF, any 2 of SPS, focal defects, or extent were sufficient to account for 43% of the variance in DCRs, and using all three texture variables plus the SPS only accounted for 44%.

In COPD, focal defects and patchiness or the SPS and patchiness accounted for 46-48% of the variance in DCRs, and using all four variables only increased the amount of variance explained to 52%.

In PE, either the SPS and focal defects or focal defects and extent accounted for 64-67% of the variance in DCRs, and

all four variables accounted for 68%. Indeed, focal defects alone accounted for 64% of the variance in the DCR for PE. Given that an NMP can be highly certain PE exists on the basis of a single, isolated defect, a simple decision rule based purely on the presence of focal defects is not implausible.

The systemic:focal relationship was reversed here. The two texture elements had correlations of .79 with PE, .70 with COPD, and .65 with CHF. This suggests that the texture elements were more important in diagnosing focal than systemic disease.

The linear models below reveal that the focal defect rating was the most important element in the prediction equation for the CHF rating, the patchy rating was the most important in predicting the COPD rating, and the focal rating was the most important in predicting the PE rating.

$$DCR_{CHF} = .08949_{PATCHY} + .56064_{FOCAL} + .11508_{EXTENT}$$

$$DCR_{CHF} = .07022_{PATCHY} + .53974_{FOCAL} + .08039_{EXTENT} + .10300_{SPS}$$

$$DCR_{COPD} = .45372_{PATCHY} + .27951_{FOCAL} + .17083_{EXTENT}$$

$$DCR_{COPD} = .43093_{PATCHY} + .24493_{FOCAL} + .10538_{EXTENT} + .16928_{SPS}$$

$$DCR_{PE} = -.11930_{PATCHY} + .65627_{FOCAL} + .29059_{EXTENT}$$

$$DCR_{PE} = -.12411_{PATCHY} + .66077_{FOCAL} + .26779_{EXTENT} + .05361_{SPS}$$

Variables entered	R	R <sup>2</sup>	St. Err	F
CHF:				
SPS	.39	.15	1.10	22.08
SPS & Extent	.44	.19	1.08	14.19
SPS & Focal	.65	.43	.91	44.97
SPS & Patchy	.44	.20	1.08	14.93
Focal	.63	.40	.93	81.86
Focal, Extent	.65	.43	.91	45.32
Focal, Patchy	.65	.43	.91	44.83
Patchy	.38	.14	1.11	20.49
Patchy, Extent	.41	.17	1.10	11.99
Extent	.55	.31	.19	60.80
SPS, Focal, Extent	.66	.44	.91	31.00
SPS, Patchy, Focal	.66	.44	.91	30.93
SPS, Patchy, Extent	.45	.20	1.08	10.28
Patchy, Focal & Extent	.66	.43	.91	30.43
Patchy, Focal, SPS, Extent.	.66	.44	.91	23.27
PE:				
SPS	.34	.11	.70	17.74
SPS & Extent	.60	.37	1.37	40.26
SPS & Patchy	.27	.07	1.66	5.33*
SPS & Focal	.80	.64	1.03	125.52
Focal	.79	.63	1.04	239.68
Focal, Extent	.82	.67	1.00	139.73
Focal, Patchy	.79	.63	1.05	119.11
Patchy	.16	.03	1.07	1.82**
Patchy, Extent	.62	.38	1.35	43.57
Extent	.60	.37	1.37	81.08
SPS, Focal, Extent	.82	.67	1.00	93.12
SPS, Patchy, Focal	.80	.64	1.03	83.76
SPS, Patchy, Extent	.62	.38	1.36	28.84
Focal, Extent, Patchy	.82	.68	.98	97.29
Focal, Extent, Patchy, SPS.	.82	.68	.98	73.22

Table 7.5(a). Correlation of diagnosis certainty ratings with texture ratings plus the SPS. Test sets for each disease rating consisted of cases which had that disease plus true normals. All F's significant at  $p \leq .00005$ , unless marked (\*\*),  $p \leq .1646$ , or (\*),  $p \leq .0059$ .

Variables entered	R	R <sup>2</sup>	St. Err	F
<b>COPD:</b>				
SPS	.49	.24	1.34	66.62
SPS & Extent	.58	.33	1.26	52.32
SPS & Patchy	.68	.46	1.13	89.74
SPS & Focal	.58	.33	1.26	51.83
Focal	.45	.20	1.38	54.12
Focal, Extent	.61	.37	1.23	61.84
Focal, Patchy	.70	.48	1.11	97.80
Patchy	.63	.40	1.19	140.58
Patchy, Extent	.66	.43	1.17	79.26
Extent	.52	.27	1.32	76.20
SPS, Focal, Extent	.64	.41	1.20	47.33
SPS, Patchy, Focal	.72	.52	1.08	73.63
SPS, Patchy, Extent	.69	.47	1.13	61.44
Patchy, Focal, Extent	.71	.50	1.09	69.89
Patchy, Focal, SPS, Extent.	.72	.52	1.08	56.34
<b>ALL CASES:</b>				
Patchy, Focal, Extent				
CHF	.45	.20	.97	24.69
COPD	.64	.41	1.12	17.52
PE	.73	.53	.97	109.54

Table 7.5(b). Correlation of diagnosis certainty ratings with texture ratings plus the SPS. Test set for COPD disease rating consisted of cases which had that disease plus true normals. Test set for all cases include all diseased cases plus all true normals. All F's significant at  $p \leq .00005$ .

#### Factor Analysis.

To further illuminate the relationship between percepts of texture and disease and actual texture, a principal components factor analysis was performed on the six rating scales using the BMDP 4M program, and 99% of the variance in factor space could be explained in three factors. Table 7.6 contains the sorted, rotated factor loadings of each rating scale.

Factor 1 loads primarily on patchiness, with large loadings of COPD and extent and a small loading of CHF. This factor may best be considered a texture factor for patchiness. Factor 2 loads mainly on PE, with approximately equal, but much smaller, loadings of extent and focal defects, and thus can be considered a PE factor. Factor 3 loads mainly on focal defects, with minor loadings of CHF and PE, and thus appears to be a texture factor for focal defects without being limited to a specific disease.

This analysis supports assumptions by NMPs and the investigator that COPD and patchiness are related concepts for NMPs, and that focality and PE are related concepts. The interpretation of loadings on CHF are less clear.

A second factor analysis was done which included the six rating scales, plus the mean and standard deviation of the SPS. Table 7.7 has the sorted, rotated factor loadings that

rating scale	factor 1	factor 2	factor 3
patchy	0.881	0.000	0.000
COPD	0.680	0.000	0.000
extent	0.532	0.440	0.000
PE	0.000	0.834	0.378
focal	0.000	0.416	0.884
CHF	0.363	0.000	0.287
variance explained:	1.670	1.065	1.029

Table 7.6 Sorted, rotated factor loadings for the six rating scales, in decreasing order by the amount of variance explained by that factor.

resulted from that analysis. In this analysis 100% of the variance in factor space could be explained in three factors.

rating scale	factor 1	factor 2	factor 3
patchy	0.834	0.000	0.000
COPD	0.816	0.000	0.000
extent	0.676	0.348	0.000
SPS	0.660	0.000	0.000
PE	0.000	0.933	0.000
focal	0.000	0.858	0.000
Standard deviation	0.000	0.000	0.845
CHF	0.360	0.000	0.678
variance explained:	2.433	1.774	1.214

Table 7.7 Sorted, rotated factor loadings for the six rating scales plus the mean and standard deviation of the SPS, in decreasing order by the amount of variance explained by that factor.

In this pattern, the first factor and second factors are very similar to those in the rating scale factor analysis. In addition, the SPS loads strongly on factor 1. The third factor, however, is no longer a focal defect factor, but is made up of the standard deviation of the SPS's and the CHF rating. The standard deviation of the SPS reflects the size of the heart impression, which is much greater in CHF than in non-CHF cases. These three factors might best be characterized as disease-specific factors, with factor 1 being COPD, factor 2 being PE, and factor 3 being CHF. Adding in a measure of the differences between views, albeit

a statistical rather than a perceptual one, clarifies the factor structure in regards to CHF.

Prediction.

Table 7.8 contains the results from ROC analyses performed in the form of the area under the distribution-free ROC curve and the associated standard error.

Each disease was treated separately. Disease sets for each AUC are composed of the cases for which that disease is positive, plus the 20 true normal cases.

The first analysis is the AUC achieved by NMPs' DCRs for each disease. The second analysis is the AUC achievable by using the SPS alone.

scores\disease	CHF	COPD	PE
(1) original DCR	.85±.14	.87±.05	.98±.02
(2) SPS	.72±.10	.72±.07	.58±.10
(3) Patchy, Focal, extent scales	.89±.10	.90±.13	.95±.04
LDFA:			
(4) SPS, s	.70±.04	.76±.04	
(5) patchy, focal scales	.86±.03	.87±.02	.97±.01

Table 7.8. Areas under the ROC curve (AUC) ± the standard error calculated from (1) original scores assigned by NMP's, (2) the SPS alone, (3) the scores derived from weighting texture ratings with the Beta weights derived from a multiple regression procedure, and two linear discriminant function analyses based on the canonical variables from: (4) the SPS and its standard deviation, and (5) the patchy and focal scales. All calculations done on a test set of one disease type plus true normals.

The SPS was significantly inferior to the DCRs ( $p=.054$ ) in discriminating COPD, and for discriminating PE ( $p\leq.00006$ ), but was not significantly different in discriminating CHF.

The third score represents the result of performing a multiple regression procedure to predict disease from texture and extent, deriving beta weights, and using those beta weight to calculate, for each image, a new "DCR" based solely on texture and extent ratings. As can be seen, this procedure results in ROCs very close to that of the original DCR.

The fourth score represents the results of calculating an AUC from the canonical variables assigned to each case as a result of a jackknifed linear discriminant analysis based on the mean and standard deviation of the SPS. The procedure from BMDP was utilized for this analysis. No discrimination was possible for PE. Although the factor analysis revealed the standard deviation of slopes to be related to the CHF factor, incorporating the standard deviation did not improve discrimination for CHF. However, incorporating the variability between views did appear to improve prediction a small amount in the case of COPD.

The fifth score represents the results of calculating an AUC from the canonical variables assigned each case as a result of the jackknifed linear discriminant analysis based on the patchy and focal rating scores. These two rating scales alone produced discrimination almost identical to that of the DCRs for all three diseases.

There were no significant differences within one disease between the AUCs generated by the original DCRs, the texture and extent scales as weighted by the physicians' linear model, and the patchy and focal scales as combined by the optimal discrimination function of linear discriminant function analysis, indicating that the physician use of texture variables is very nearly optimal for the discrimination of disease.

## CHAPTER 8

## Discussion

This investigation examined relationships between statistical and perceived texture in scintigraphic lung scans, and the relative abilities of the statistical measure of texture and of the physicians to detect disease. The results of the study provide the basis for the following discussion, conclusions, and recommendations.

Discussion of Results and Conclusions

In the introduction, three questions were posed. These will be addressed in the order in which they were presented.

The first question asked was "What is the relationship between texture as perceived by nuclear medicine physicians and texture as described by the slope of the power spectrum?" This question was of interest because the SPS is a measure related to fractal dimension; of concern is the ability of NMPs to perceive texture relevant to the fractal dimension.

Previous studies, while finding high correlations between fractal dimension and such perceptual dimensions as roughness and complexity, have used fractal simulations of lines and surfaces. Such images are relatively easy for the perceptual system to decipher. In this study, however, the images used are two-dimensional projections of a three dimensional stochastic fractal, and their interpretation is

based on sophisticated expert knowledge of all the systems involved, physiological and technical. Moreover, the measures of texture utilized were not simple, as, for instance, rating the "roughness" of an image is. Rather, they require the NMP to extrapolate from the image to the 3-D structure of the lungs, and evaluate the relative grey levels present in the image in terms of what it reflects about the physical texture, function and shape of the lung.

It was by no means certain that the fractal dimension would capture any of this complex decision-making process, but as the results show, the SPS is significantly related to the measures of texture used by the NMPs.

The size of the relationship between texture and SPS depends very much on the disease present in the lung, with the systemic:focal dimension influencing the size of the correlation. In the systemic disease of CHF the patchiness and focality certainty ratings explain about 30% of the variance in the SPS; in the focal disease, PE, about 9%; and in the disease which was both focal and systemic, COPD, about 23%.

In addition, the extensiveness of the disease in the lung within each disease type influenced the correlation of texture to SPS. The more widespread the disease was perceived to be, the higher the correlation between texture

ratings and SPS. Since the SPS measures the global texture of the image, while the NMPs are free to restrict (and by verbal report do restrict) their texture analysis to those parts of the lung which appear diseased, such a pattern of relationships seems reasonable.

The second question asked was "To what extent is texture relevant to diagnosis?"

Within one disease type, perceived texture explains a large proportion of the variance in diagnostic ratings. The two texture ratings predicted from 43% (CHF) up to 63% (PE).

Correlation between DCRs and texture ratings was much lower where all disease types were included in the analysis (20-53%). This suggests that the texture ratings are not used consistently between disease types.

The SPS does not account for as much of the DCRs as the perceived texture. The SPS accounts for no more than 24% of the variance of any DCR.

The third question was "Which measure better captures the texture relevant to discrimination of disease: perceived texture ratings, or the SPS?". The judgments of the NMPs were superior to the SPS, both in terms of their nosological judgments (the DCRs) and the contributory percepts (the texture ratings). This was true whether the mean SPS alone was used as the discriminating variable, or whether a linear

discriminant analysis which included both the mean SPS and the SPSs' standard deviation was used. This superiority of the NMP may be because the SPS examines only texture, and the NMPs examine texture plus other facets of the image which are relevant to diagnosis.

The SPS, however, was used as a texture metric because it is a measure linearly related to the fractal dimension of an image. Images which vary widely in appearance may have identical dimensionality. Although lung scans are expected to have similar dimensionality based on similar physiological structures, it is possible that disease entities may affect the dimensionality to the same degree and in the same direction while influencing the physiological structure into diverging patterns. To completely describe a fractal, one must describe both the initiator and the generator, a procedure not yet possible from lung scans. Therefore the SPS is, necessarily, an incomplete descriptor of a lung scan, and it appears that much of the information needed to create an accurate diagnosis is in elements of the image not captured by the SPS.

#### Recommendations.

This study serves two purposes: to evaluate the SPS as a predictor of disease, and to investigate the perception of fractal dimension in complex stochastic fractals.

As a predictor of disease, the SPS proved itself inferior to the decision-making process of nuclear medicine physicians. The diagnosis of a disease from scintigraphs is a complex process, requiring extensive familiarity with the technology involved, the anatomy and function of the organs imaged, and the pathologies possible in the image. Although NMPs attribute much of their decision-making process to the analysis of texture, this study reveals that the process is not so simple as they describe. Although it is a comprehensive statistical measure of texture, the SPS did not explain the majority of their diagnostic certainty, nor is it an adequate predictor of disease. The superior ability of the NMPs to detect disease demonstrates the incorporation of information not reflected in the SPS.

However, the SPS is not irrelevant to the decision process. The SPS was significantly correlated to all perceptual measures, and predicted disease at a level far above chance. It may prove a valuable component of future machine-based diagnostic algorithms.

One limitation of the SPS as a predictor of disease appears to be that it provides an overall measure of texture rather than one localized to the diseased part of the image. Analysis of smaller units in the image may provide better discrimination, and should be investigated.

The SPS was found to explain some variance in the perceptual measures of texture, but was not used consistently across disease types. The issue of how it was used differently was not addressed in this study, and remains for future research to investigate.

As an explanation of texture perception, the SPS might prove useful for less complex images, especially ones which do not require expert knowledge. Its ability to predict perceptions of roughness and smoothness, among other textural perceptions, should be investigated.

In addition, future research into the evaluation of the perception of fractal dimension would benefit from taking a path between the overly simplistic fractals of previous work and the stochastic nature of the fractals of the current work. Various fractal-generation programs exist (for example, those in Stevens, 1989) which could be used to create fractals of high levels of complexity and definable dimensionality. Software could be used to generate digital images of landscapes, that, while seeming to be naturalistic, could be precisely specified mathematically (Feder). Such images might tell us a great deal about our perception of texture in the world we live in.

## APPENDIX A: BETA WEIGHTS

This appendix contains beta weights for all multiple correlations tabled in Chapter 7.

Table 7.4. Correlations of texture variables with the SPS.

analysis	PATCHY	FOCAL	EXTENT
CHF:			
Patchy & Focal	.43552	.20649	
Patchy & Extent	.22546		.39226
Extent & Focal		.21343	.49237
Patchy, Focal & Extent	.17402	.18796	.37553
COPD:			
Patchy & Focal	.35174	.24853	
Patchy & Extent	.17198		.41991
Extent & Focal		.22150	.46187
Patchy, Focal & Extent	.13462	.20429	.38664
PE:			
Patchy & Focal	.24145	.12194	
Patchy & Extent	.09414		.37614
Extent & Focal		-.08929	.46986
Patchy, Focal & Extent	.08974	-.08389	.42532
ALL CASES:			
Patchy & Focal	.32791	.14148	
Patchy & Extent	.16226		.40203
Extent & Focal		.07236	.45873
Patchy, Focal & Extent	.16026	.06842	.38590

Table 7.5. Correlations of diagnostic certainty ratings with texture ratings and the SPS.

analysis	Patchy	Focal	Extent	SPS
CHF:				
SPS & patchy	.24431			.26901
SPS & focal		.56722		.17092
SPS & extent			.22772	.26858
patchy & focal	.16705	.57170		
patchy & extent	.23107		.20872	
focal & extent		.57338	.17417	
SPS, patchy, focal	.11659	.54303		.12186
SPS, patchy, extent	.17397		.11995	.23953
SPS, focal, extent		.54756	.12239	.11249
patchy, focal, extent	.08949	.56064	.11508	
patchy, focal, extent, SPS	.07022	.53974	.08039	.10300
COPD:				
SPS & patchy	.51853			.27235
SPS & focal		.32111		.37938
SPS & extent			.35712	.30434
patchy & focal	.54965	.29906		
patchy & extent	.50484		.21635	
focal & extent		.33752	.42438	
SPS, patchy, focal	.47821	.24858		.20311
SPS, patchy, extent	.46473		.11843	.23319
SPS, focal, extent		.28811	.32134	.22309
patchy, focal, extent	.45372	.27951	.17083	
patchy, focal, extent, SPS	.43093	.24493	.10538	.16928
PE:				
SPS & patchy	.09929			.22158
SPS & focal		.77374		.11298
SPS & extent			.60696	-.00653
patchy & focal	-.01565	.79690		
patchy & extent	-.15368		.67530	
focal & extent		.66345	.23137	
SPS, patchy, focal	-.04553	.78181		.12378
SPS, patchy, extent	-.15437		.67254	.00733
SPS, focal, extent		.66729	.21119	.04295
patchy, focal, extent	-.11930	.65627	.29059	
patchy, focal, extent, SPS	-.12411	.66077	.26779	.05361

**APPENDIX B: ABBREVIATIONS**

To assist the reader of this study, all abbreviations used are defined below.

AUC: Area under the Receiving Operating Characteristic Curve  
c/cm: Cycles per centimeter  
c/deg: Cycles per degree of visual angle  
CHF: Congestive Heart Failure  
COPD: Chronic Obstructive Pulmonary Disease  
DCR: Diagnostic Certainty Rating  
FFT: Fast Fourier Transform  
NMP: Nuclear Medicine Physician  
PE: Pulmonary Embolism  
ROC: Receiver Operating Characteristic  
SDT: Signal Detection Theory  
SPS: the Slope of the Power Spectrum

## REFERENCES

- Bamber, D. (1975). The area above the ordinal dominance graph and the area below the receiver operating characteristic graph. Journal of Mathematical Psychology, 12, 387-415.
- Blessner, B. & Ozonoff, D. (1972). A model for the radiologic process. Radiology, 103, 515-521.
- Campbell, F. W. & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings, Journal of Physiology (London), 197, 551-566.
- Cargill, E. B. (1989). A mathematical liver model and its application to system optimization and texture analysis. (Doctoral Dissertation, University of Arizona, 1988). Dissertation Abstracts International, 51/01B, 467.
- Cutting, J. E. & Garvin, J. J. (1987). Fractal curves and complexity. Perception and psychophysics, 42(4), 365-370.
- DeNardo, G. L. (1977). Lung imaging. in P. Matin (Ed.), Handbook of Clinical Nuclear Medicine. Flushing, NY: Medical Examination Publishing Company, Inc.
- DeNardo, G. L. & DeNardo, S. J. (1984) The Lungs. in L. M. Freeman, (Ed.), Freeman and Johnson's clinical radionuclide imaging (Vol. 2) (3rd ed.). Philadelphia, PA: W.B. Saunders Co.
- Dorfman, D. D., Alf, E. Jr. (1969). Maximum likelihood estimation of parameters of signal detection theory and determination of confidence intervals: Rating method data. Journal of Mathematical Psychology, 6, 487-496.
- Feder, J. (1988). Fractals. New York: Plenum Press.
- Friel, J. P. (Ed.) (1974). Dorland's illustrated medical dictionary (25th ed.). Philadelphia: W. B. Saunders Co.
- Goldberger, A. L., Bhargava, V., West, B. J. & Mandell, A. J. (1985). On a mechanism of cardiac electrical stability. Biophysical Journal, 48, 525-528.

- Goldberger, A. L., Rigney, D. R., & West, B. J. (1990). Chaos and fractals in human physiology. Scientific American, Feb., 43-50.
- Goldberger, A. L. & West, G. J. (1987). Fractals in physiology and medicine. The Yale Journal of Biology and Medicine, 60, 421-435.
- Goris, M. L. (1985). Sensitivity and specificity of common scintigraphic procedures. Chicago: Year Book Medical Publishers, Inc.
- Gray, H. L. & Schucany, W. R. (1972). The Generalized Jackknife Statistic. New York: Marcel Dekker.
- Grossberg, S. & Mingolla, E. (1985). Neural dynamics of perceptual grouping: Textures, boundaries and emergent segmentations. Perception and psychophysics, 38, 141-171.
- Hanley, J. A. & McNeil, B. J. (1982). The meaning and use of the area under a receiving operating characteristic (ROC) curve. Radiology, 143, 29-36.
- Harvey, L. O. & Gervais, M. J. (1978). Visual texture perception and Fourier analysis. Perception and psychophysics, 24, 534-542.
- Hayes, A. (1988). Identification of two-tone images; some implications for high- and low-spatial-frequency processes in human vision. Perception, 17, 429-436.
- Ker, M., Seeley, G. W., Stempski, M. O., & Patton, D. (1988). A protocol for verifying truth of diagnosis: Preliminary report. Investigative Radiology, 23, 485-487.
- Kim, E. E. & Haynie, T. P. (1987). Nuclear diagnostic imaging: Practical clinical applications. New York, NY: MacMillan Publishing Co.
- Kundel, H. L. & Revesz, G. (1974). The evaluation of radiographic techniques by observer tests: Problems, pitfalls, and procedures. Investigative Radiology, 9, 166-173.

- Lefevre, J. (1983). Teleonomical optimization of a fractal model of the pulmonary arterial bed. Journal of Theoretical Biology, 102, 225-248.
- Macmillan, N. A. & Kaplan, H. L. (1985). Detection theory analysis of group data: Estimating sensitivity from average hit and false-alarm rates. Psychological Bulletin, 98, 185-199.
- Mandelbrot, B. B. (1977). The Fractal Geometry of Nature. San Francisco, CA: W. H. Freeman and Company.
- Massof, R. W. & Emmel, T. C. (1986). Criterion-free parameter-free distribution-independent index of diagnostic test performance. Applied Optics, 26, 1395-1408.
- Mohsenifar, Z., Amin, D. K., & Shah, P. K. (1989). Regional distribution of lung perfusion and ventilation in patients with chronic congestive heart failure and its relationship to cardiopulmonary hemodynamics. American Heart Journal, 117, 887-891.
- McClish, D. K. (1987). Comparing the areas under more than two independent ROC curves. Medical Decision Making, 7, 149-155.
- McNeil, B. J. (1980). Ventilation-perfusion studies and the diagnosis of pulmonary embolism: Concise communication. Journal of Nuclear Medicine, 21, 319-323.
- McNeil, B. J. & Hanley, J. A. (1984). Statistical approaches to the analysis of the receiver operating characteristic (ROC) curves. Medical Decision Making, 4, 137-149.
- Metz, C. E. (1978). Basic principles of ROC analysis. Seminars in Nuclear Medicine, 8, 283-298.
- Muehllehner, G. (1980). Scintillation camera collimators, in Nudelman, S. & Patton, D. D. (Eds.), Imaging for medicine, (Vol. 1). NY: Plenum Press.
- Murray, J. F. (1986). The normal lung. Philadelphia: W.B. Saunders Co.

- Murrell, G. A. (1977). Combination of evidence in perceptual judgment. In M. F. Kaplan, & S. Schwartz (Eds.), Human judgment and decision processes in applied settings (pp. 169-201). NY: Academic Press.
- Pentland, A. P. (1984). Fractal-based description of natural scenes. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-6, 661-674.
- Peterson, W. W., Birdsall, T. G., Fox, W. C. (1954). The theory of signal detectability. Transactions of the IRE Professional Group on Information Theory, 4, 171-212. (Reprinted in R. D. Luce, R. R. Bush, & E. Galanter, (Eds.), (1963). Readings in mathematical psychology, 167-211. New York: Wiley.
- Pitz, G. F. & Sachs, N. J. (1984). Judgment and decision: theory and application. Annual Review of Psychology, 35, 139-163.
- Quinn, P. C. (1985). Suprathreshold contrast perception as a function of spatial frequency. Perception and psychophysics, 38, 408-414.
- Sagi, D. (1988). The combination of spatial frequency and orientation is effortlessly perceived. Perception and psychophysics, 43, 601-603.
- Sagi, D. & Hochstein, S. (1984). The contrast dependence of partial frequency channel interactions. Vision Research, 24, 1357-1365.
- Schwartz, S. & Griffin, T. (1986). Medical thinking: The psychology of medical judgment and decision making. New York, NY: Springer-Verlag New York, Inc.
- Sekuler, R., Wilson, H. R., & Owsley, C. (1984). Structural modelling of spatial vision. Vision Research, 24, 689-700.
- Spencer, R. O. (1980). Nuclear medicine: Focus on clinical diagnosis, (2nd ed.). Garden City, NY: Medical Examination Publishing Company.
- Stevens, R. T. (1989). Fractal Programming in C. Redwood City, CA: M & T Books.

- Susskind, H. & Rezak, I. (1984). Lung physiology and anatomy, in Atkins, H. L. (Ed.), Pulmonary nuclear medicine: Techniques in diagnosis of lung disease. NY: Marcel Dekker, Inc.
- Swets, J. A. (1986). Indices of discrimination or diagnostic accuracy: Their ROCs and implied models. Psychological Bulletin, 99, 100-117.
- Swets, J. A. (1986). Form of empirical ROCs in discrimination and diagnostic tasks: Implications for theory and measurement of performance. Psychological Bulletin, 99, 181-198.
- Sutter, A. & Beck, J. (1989). Contrast and spatial variables in texture segregation: Testing a simple spatial-frequency channels model. Perception and psychophysics, 46, 312-332.
- Tanaka, T., Katou, K., Kimata, S., Hosada S., & Hirose K. (1989). Severity of cardiac failure from the standpoint of pulmonary circulation: Studies centered on the distribution of pulmonary perfusion. Japanese Circulation Journal, 53, 155-164.