STATISTICAL ANALYSIS
OF
SPECT NEUROIMAGES
IN
ALZHEIMER'S DISEASE

by

Stephen James McCrory

Thesis submitted for the degree
of Master of Science
in the University of Glasgow

December 1992

ACKNOWLEDGEMENTS

I am grateful to my supervisor Dr. Ian Ford in the Department of Statistics and members of the Alzheimer project at the University of Glasgow for their help throughout the period of this research.

TABLE OF CONTENTS

SUMMARY

We have considered a variety of problems in the statistical analysis of quantitative data extracted from Single Photon Emission Computed Tomography (SPECT) and Positron Emission Tomography (PET) neuroimages. Many of the questions of interest to investigators in this area of research were described and illustrated with analysis of some SPECT datasets.

In chapter one we introduced the technique of tomograhic imaging, describing a common approach to extracting quantitative data from images and gave some background to Alzheimer's disease. In chapter two we described a major research study into Alzheimer's disease, from which most of the datasets used in this thesis were obtained.

In chapter three we identified the broad categories of statistical problems into within and between group analysis of regional mean patterns and the interrelationships among regions. A SPECT dataset, consisting of Alzheimer and normal control subjects, was used to illustrate the use of univariate methods to study these problems. We saw from these analyses that it was difficult to extract clear biological interpretations with this approach due to certain features in the data extracted from the images. In particular, the presence of substantial random variation between subject data vectors meant that meaningful analysis could only be carried out after adjusting the regional data - to remove the between

subject variation - prior to the analysis. Different methods of adjustment were seen to give different results here. Although not particularly evident from these data, another feature of typical imaging datasets was the large number of regions to be analysed.

In chapter four we looked at the application of univariate and multivariate ANOVA type methods to compare regional mean profiles between groups and illustrated some approaches to follow-up analysis. Assumptions underlying these techniques, including normality and equality of covariance matrices were assessed as was the choice of scale for the analysis. The assumption of multivariate normality was reasonable on the square root scale in both groups, although equality of group covariance matrices was very strongly rejected. Even though many of the assumptions in the RM ANOVA may be violated for these datasets, the fact that global tests can be performed, even when the number of regions p exceeds the numbers of subject n, will make this the most viable approach. Adjusted F-tests will be appropriate in such circumstances.

In chapter five, we looked at some approaches to investigating inter-relationships among regions. The most common approach here is to use simple correlation analysis among regions after adjusting data vectors for the subject effect. As in the analysis of means, the results will be strongly influenced by the form of the adjustment. Ford (1986) has shown that inferences from the results of correlation analysis are made difficult;

with adjustment in the data resulting in confounding of parameters in a model of the correlation structure. Even so, between group comparisons may still provide valuable insight into a disease process. A testing scheme gave tentative evidence of differences in the correlation structure between the normal and Alzheimer groups. Multivariate exploratory techniques where used to study the interregional correlation structures. Principal components analysis demonstrated that just a few patterns accounted for most of the variation among subjects in each of the groups and that bilateral pairs of regions were very strongly correlated. Further canonical correlation analysis of the data suggested that regional profiles may be summarised into hemispheric sums and differences separately without too much loss of information. In studies with several regions being studied this would be a useful reduction of dimensions. Multi-Dimensional scaling highlighted a number of other features in our data including the measurement difficulties with smallish regions such as the basal-ganglia and some evidence of a spatial relationship between regional data. Formal analysis of the covariance structures using the spatial correlation model of Worsley et. al. (1992) gave some evidence for this feature in SPECT data, albeit using an estimated distance matrix.

The analysis of repeat scan data, as arise in longitudinal studies, was considered in Chapter six. The univariate RM ANOVA approach of chapter 4 was used to analyse a dataset from the Alzheimer study. Evidence was

provided of significant changes over time in the average profiles.

In chapter seven we looked at approaches to correlating imaging data with other datasets studied in Alzheimer research including neuropsychological and post mortem plaque density data. As with simple correlation type analyses in interregional studies, the SPECT data had to be adjusted prior to calculation of correlation coefficients. As before this meant that different results were given with different forms of adjustment which did not necessarily give any clear biological insight into the data. Multivariate regression, canonical correlation analysis and an application of PC analysis were used to illustrate some alternative approaches. The within-subject approach taken to compare SPECT profiles with PM plaque density was a useful illustration of the sort of 'clever' approaches employed by researchers in brain research to analysing their data. The fact that it was statistically incorrect highlights a feature of much of the analysis of imaging datasets; lots of data amassed without any great insight into appropriate methods of analysis.

Further topics of interest were considered in chapter eight, including the use of these data in discrimination type analysis and problems of measurement. We also introduced the brain activation paradigm which is being increasingly used to study many of the problems covered in the thesis.

Finally, in chapter nine, we considered any conclusions that could be drawn and suggested areas of analysis requiring further work.

CHAPTER 1

THE BRAIN: STUDY OF A COMPLEX HETEROGENEOUS ORGAN!

1.1 INTRODUCTION

In the last two decades, advances in medical scanning technology have led to the development of a number of imaging modalities capable of providing structural and/or functional insight into the living human body, and in particular the human brain. The main functional modalities of importance at the present time include positron emission tomography (PET) and single photon emission tomography (SPECT). These modalities and the various imaging methods used make it possible to obtain a picture of activity in the living brain from which large amounts of quantitative and qualitative data can be extracted, relating to a variety of biological processes, including blood flow, metabolic & physiological activity and brain receptor status.

Consequently, both PET and SPECT are being utilised in an increasingly wide variety of clinical applications as well as in fundamental scientific studies of the brain and brain disorders. Clinical applications are found in cardiology, oncology, neurosurgery and neurology where regions of low blood flow may be important indicators of tumour growth or ischaemic damage resulting from stroke and/or head injury. Experimental applications are found in the study of normal metabolic activity and changes rising from the impact of different diseases e.g.

Alzheimer's Disease, Parkinson's Disease, AIDS and cerebrovascular disease. Other applications of functional imaging include the study of higher mental function through brain activation through stimulation or after pharmacokinetic intervention.

A number of issues arise in handling the statistical analysis of datasets generated from these imaging studies. In this thesis we will address some of these issues, with reference to some of the important questions of interest to investigators and provide examples of the sorts of statisical approaches which are and can be taken to analyse images. Various SPECT datasets, obtained from a research study of Alzheimer's Disease, are used for this purpose.

1.2 TOMOGRAPHIC IMAGING OF THE LIVING BRAIN

Tomography is a branch of nuclear radiology that is concerned with the construction of an image of physiological activity or structural anatomy through a living body without the need for invasive surgery. Consequently, tomographic imaging has much appeal for studying a complex organ such as the human brain. The genesis of the current techniques used in tomographic imaging are based on a geometrical result, originally derived by the Austrian mathematician J. Radon. His thesis was that any function of two variables could be constructed if the infinite set of all its line integrals were known. A simple numerical example of the concept is

given in the next section.

For the purposes of medical imaging, the study of cerebral function is mediated through the use of radioactive tracer isotopes and radiation detection technology. The radioactive isotope is biochemically engineered to bind to a compound which is introduced into the body usually through the bloodstream. The blood acts as a carrier (of the isotope) to the organ or site of interest. An example is the fluorodeoxyglucose technique used in PET studies. The decay of the isotope at this site can be traced using sophisticated camera equipment and mathematical/statistical techniques used to estimate the levels of acivity e.g. blood flow at this site. In imaging the brain, a ring of detectors is placed around the subject's head providing full 360^{o} coverage.

The radiation emissions are counted by each detector and processed through a statistical reconstruction algorithm to produce 2- and sometimes 3-dimensional image datasets (depending on the imaging system and reconstruction algorithm). These images are in the form of a matrix of pixel or voxel counts involving many thousands of data values; again, the number will depend on the resolution of the system used. In order to summarise these data, and to aid clinical evaluation, the image is often thresholded to form a contour map of radiation intensity. By implication the intensity of radiation activity reflects the level of activity of the variable under study e.g. blood flow.

An example of an image obtained using SPECT is shown

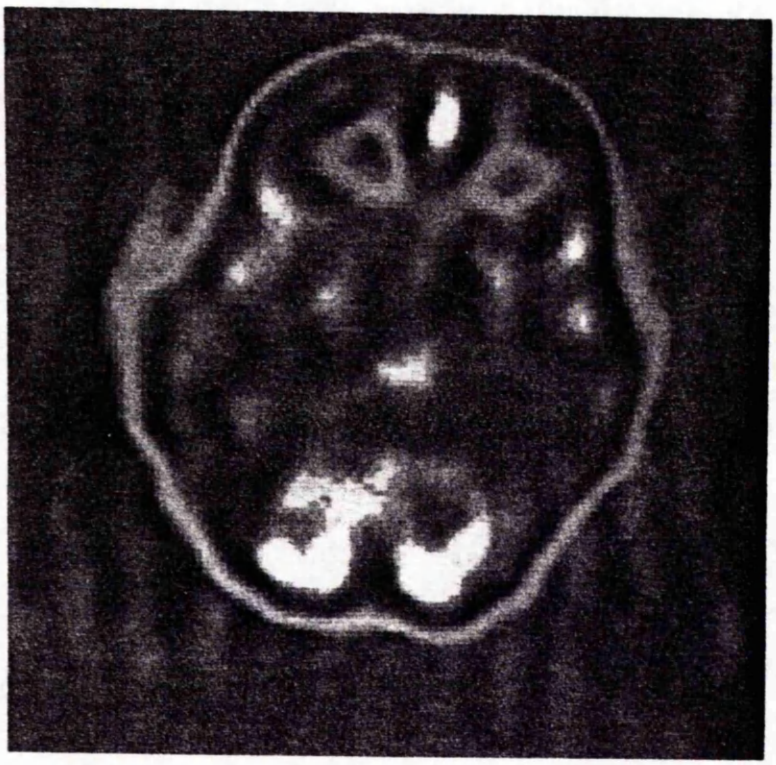


Figure 1.1 SPECT image of a normal subject.

in Figure 1.1. This shows a transverse section of radiation activity through the brain of a healthy elderly man. The bright (orange and white) areas reflect relative high radioactivity and hence high blood flow, in contrast to the darker (blue) areas - reflecting relatively lower blood flow. If studied closely, individual pixel boxes can be seen in this image.

Although based on the same underlying principles, PET and SPECT differ in a number of ways. For instance, each is based on different geometries of radiation detection. Briefly, SPECT involves detection of a single photon emission while PET involves coincident detection of a pair of photons emitted from a point source and at 180^{O} from each other. They also differ in the isotopes used. With PET a variety of positron emitting isotopes are available, allowing the mapping of a number of metabolic, physiological and receptor variables. An additional feature of PET imaging is that it is possible to model the tracer activity in terms of more immediately appreciable biological units, using compartmental modelling techniques (Phelps et al, 1979). This is because of the short half-lives of the isotopes used. By contrast, there are only a few isotopes with much longer half-lives available for SPECT imaging, although a number of new ones are under development (Ell, 1990).

1.3 IMAGING DATA, RECONSTRUCTION AND ROI EXTRACTION

The data used for statistical analysis is the result

of a two phase process. The first phase involves image reconstruction; the second, image analysis.

Because of data limitations and methodological difficulties in reconstruction, the reconstructed image is only an estimate of the true picture of activity. Various sources of methodological and random error will affect the precision of a given image. In order to appreciate some of the statistical problems involved we consider the following simple example. Suppose we have a 2x2 array of 4 pixels (Figure 1.2a) for which we wish to reconstruct the pixel counts. If the row, column and diagonal totals are known, representing the infinite set of line integral data, then a uniquely identifiable solution is available (Figure 1.2b).

Figure 1.2a & b. Reconstruction in a 2x2 array of 4 pixels.

However, if the diagonal totals are omitted, a number of solutions are possible. Alternatively, if independent noise were added to these marginal totals, no unique solution would be available. In these circumstances, reconstruction becomes a problem of statistical estimation. In real problems reconstruction will involve problems of singularity and random noise. These, and other statistical issues in image reconstruction have received a great deal of attention in a broad range of journals, with some of the statistical issues being considered by Vardi et al (1985) and by Silverman et al, (1990).

In clinical applications, second phase analysis will often involve the qualitative study of an image for evidence of tumour growth, infarcts or low blood flow. However, in experimental studies it will be more appropriate to use the quantitative data to address biological questions of interest. Although some studies have carried out statistical analysis using pixel data (Fox et al, 1988; Friston et al, 1989, 1990, 1991), the dimensionality of the datasets make this a particularly difficult proposition. A modern PET system will facilitate the imaging of around 8-14 slices/subject, each slice generating around 128x128 data values/image slice. Since brains differ in size and shape, the discretised pixel image will not facilitate direct comparison of individual pixel data between subjects. For this reason, the standard approach has been to reexpress the pixel image in terms of larger anatomical regions of

interest (ROI) - notionally, the smallest experimental units for comparison among subjects. These ROIs are identified on the image using a standardised reference brain atlas, sometimes with the assistance of matched CT or MRI images (Evans, 1991). Several, say p, ROIs may be identified in this way.

AUTHORS	Imaging Modality	Number of Regions	Control	Disease
Clark et al, 1985	PET	53	15	-
Haxby et al, 1985	PET	59	26	10
Horwitz et al, 1987	PET	59	21	21
Worsley et al, 1991	PET	30	20	-
Perani et al, 1988	SPECT	48	34	-
Burns et al, 1989	SPECT	10	6	20
Montaldi et al, 1990	SPECT	14	10	26

Table 1.1 **Dimensionality in some imaging studies.**

Data is extracted from these ROI and usually summarised as a single standardised value for comparison among subjects in subsequent statistical analysis. In SPECT, we typically use mean counts/pixel. We thus obtain a vector $\underline{X}_{ij}$,

$$\underline{X}_{ij} = (X_{ij1}, X_{ij2}, \ldots\ldots\ldots\ldots\ldots, X_{ijp}) ,$$

of p regional mean counts/pixel for each subject i, in disease or treatment group j.

Despite this reduction in data volume, studies can still result in many regions being measured in a relatively small group of subjects. Some datasets, found in the neuroimaging literature, are described in Table 1.1 in terms of their dimensionality. One simlpe reason for this situation is the basic cost of experimentation, which makes it important for investigators to extract as much data as possible from each subject's image.

1.4 ALZHEIMER'S DISEASE

Alzheimer's Disease, or as it is more correctly referred to, Dementia of The Alzheimer Type (DAT), is an illness of the brain which occurs in the elderly population. It is named after the German neurologist/ physician Alois Alzheimer who noted, in 1904, a characteristic form of cell degeneration from post-mortem study of brain tissue in people who had shown particular dementing symptoms in life. These included the presence of fibrillary tangles and granular plaques.

In life DAT is characterised by progressive degeneration of mental, cognitive and behavioural function. However, clinical DAT can be difficult to diagnose as its symptoms are similar in some respects to

other dementias. An example is the similarity of between DAT and multi-infarct dementia (which can result from a stroke). In this case X-ray computed tomography is often used. Indeed clinical diagnostic criteria do exist, but they can appear quite vague:

> `Dementia is the decline of memory and other cognitive function in comparison with the patient's previous level of function as determined by a history of decline in performance and by abnormalities noted from clinical examination and neuropsychological tests. A diagnosis cannot be made when consciousness is impaired by delirium, drowsiness, stupor or coma or when other clinical abnormalities prevent adequate evaluation of mental status...'.
>
> US Dept of Health and Human Services, 1984.

Consequently, definitive diagnosis is generally made only after post-mortem examination of brain tissue for signs of tangles and plaques. Even so the criteria used often differ from laboratory to laboratory (Wiesziewsky, 1988).

Because of such diagnostic difficulties, reliable epidemiological study of this disease is also difficult (Schoenberg, 1986). Various estimates have been given for the incidence of DAT in the elderly population. See Table 1.2. Mortality estimates are also suspect since clinical DAT is not normally given as the primary cause of death. Instead, DAT is commonly given as a secondary cause

after, for example, Bronchopneumonia. Molsa et al (1982) reported evidence of 70% of the patients in their study of DAT as dying of bronchopneumonia. Katzman (1976) estimated DAT to be the fourth highest cause of mortality in the US.

	US	UK
>65 years	10%	5%
>80 years	47%	15%

Table 1.2 Age Related Incidence of DAT in UK and US.

Despite the amount of research carried out into this disease, it is surprising that few risk factors have been identified. Age and country are cited as contributing risk factors. No single cause has been identified either. One source is the genetic link (Jacob Huff et al, 1988; Martin et al, 1990). In the US the genetic link has been variously cited as accounting for 10-30% of cases. Environmental neuro-toxins have also been blamed - a popular example being aluminium. In past years this received a lot of media attention, notably through the probably incorrect idea that it was ingested through food cooked on everyday household utensils e.g. aluminium

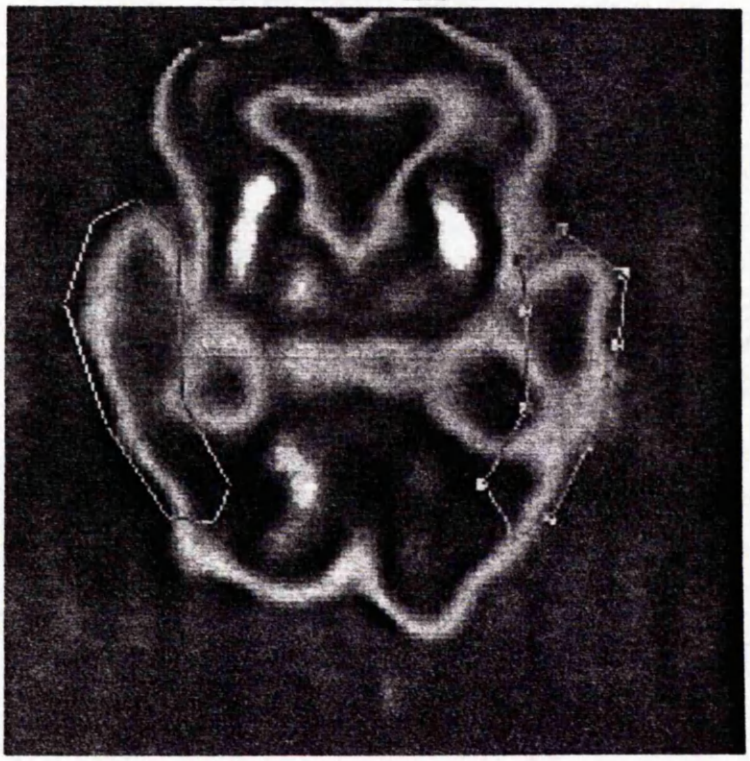


Figure 1.1 SPECT image of a DAT subject.

frying pans. A recent update of this, is that the aluminium is absorbed into the brain through nasal inhalation (Perl, 1989). Needless to say, no cures exist.

The consequence of DAT, and other mental diseases, is the devastating impact they have had on the community and on health care resources. For family, friends and other caregivers, looking after a DAT sufferer can be a tremendous psychological strain. It has even been suggested that caregivers become more vulnerable to health problems themselves as a consequence. The burden on health care resources has also been great with the need for dedicated day centres and the use of ambulances to ferry cases to and from home.

As relatively little is known about the disease functional tomographic imaging is seen as a potentially valuable tool in the study of DAT and other mental illnesses. An example of a SPECT image of a subject with a clinical diagnosis of DAT is given in figure 1.3. This shows an image for an elderly woman, exhibiting the often reported classical picture of decreased cerebral blood flow in the posterior-temporal region (both left and right of the image). Comparison of this image with the relatively high perfusion of blood flow in the corresponding area in the normal picture in figure 1.1, shows the qualitative insight this technique offers.

1.5 SCOPE OF THESIS

Functional tomographic imaging is undoubtedly an

exciting technique for the study of complex organs such as the living human brain. The aim of this thesis is to introduce the reader to this area of medical study, to identify the main questions of interest to investigators in functional tomographic brain studies and to illustrate the statistical issues through analysis of typical datasets extracted from images.

In chapter two an overview is given of the Glasgow Alzheimer project and of the different types of data collected. Chapter three introduces some of the general statistical issues in the analysis of regional means using traditional univariate approaches taken. In chapter four we look at the application of Repeated Measures ANOVA techniques to this dataset and in chapter five look at methods of investigating the inter-relationships among regions. In chapter six we consider particular problems in analysing repeated scan data, as in longitudinal studies. In chapter seven we consider approaches for correlating imaging data with other datasets observed in the Alzheimer project. In chapter eight we take a brief look at the use of imaging data in discrimination problems and some of the sources of measurement error associated with imaging datasets.

CHAPTER 2

THE ALZHEIMER PROJECT

In 1986 the Wellcome Neuroscience Group was formed at the University of Glasgow. The aim was to collect together a multidisciplinary research team to provide the basis for a long term study of Dementia of the Alzheimer Type (DAT) in the Glasgow area. The Alzheimer Project was funded principally from a grant awarded by The Wellcome Trust.

2.1 STUDY DESIGN

The core of the research programme is a 10 year longitudinal follow up study of subjects suffering from DAT. This involves six-monthly follow up evaluation of subjects until endpoint (i.e. death, withdrawal or project completion) occurs. Other related studies are also carried out, ultimately feeding into the core study. For example, small numbers of subjects in other dementia groups are studied. These include cases of a. multi-infarct dementia (MID) b. DAT & MID together and c. dementia secondary to other causes. A control group of age matched elderly normal subjects are also studied.

2.2 SUBJECT RECRUITMENT

Subjects are recruited from a variety of sources in the Glasgow area, as summarised in figure 2.1. In the

early stages of the project, DAT subjects were recruited almost entirely from long stay psycho-geriatric wards in Glasgow hospitals (e.g. Gartnavel Royal Hospital), with some others coming from old people's homes or as doctors' referrals. These were almost all cases with a well advanced, in terms of severity, form of the disease. In order to boost numbers and more importantly to obtain early stage dements, a memory clinic was set up. Advertising in local newspapers has been used for this purpose. This has proven to be a useful screening method for recruiting the sorts of dementing subjects required. Normal controls have come mainly from old peoples homes, although some are relatives of subjects in the DAT group.

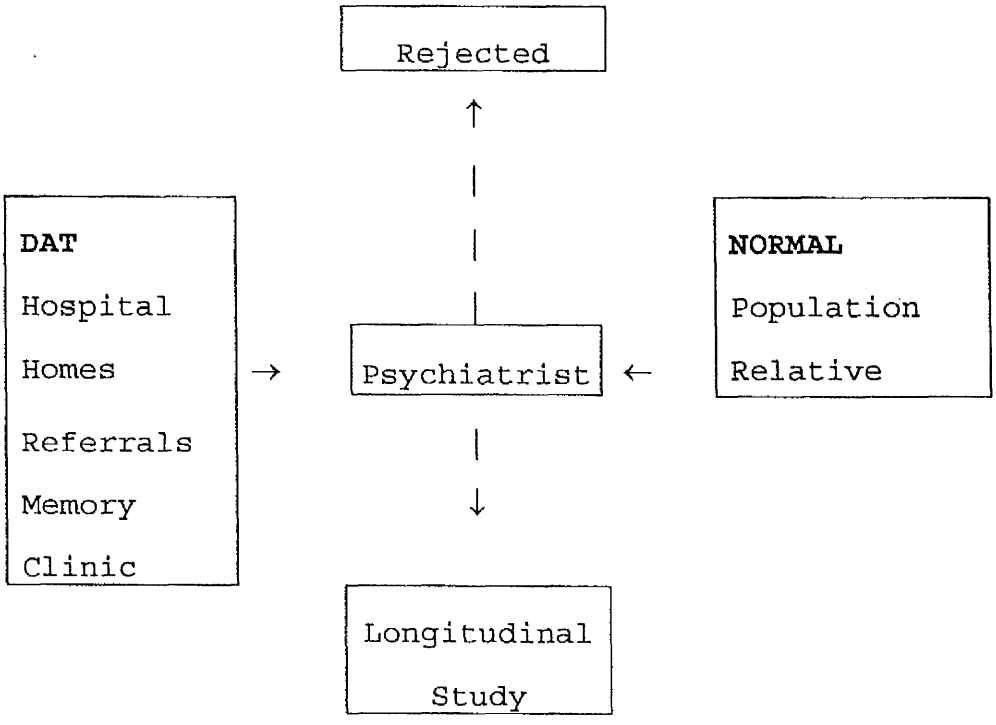


Figure 2.1 Subject Recruitment.

Subjects with a putative diagnosis of DAT undergo an extensive psychiatric assessment in addition to physical examination, laboratory testing of blood samples, subject interview and X-ray computed tomography. Only if diagnosis is confirmed, are subjects finally recruited into the longitudinal study.

2.3 STUDY DATA

A useful classification of the data collected from these subjects is (1) Diagnostic (as described above), (2) Neuropsychological test scores, (3) Functional Imaging using SPECT and (4) Post-Mortem quantitative study of Brain Tissue. These are described in more detail later.

STUDY GROUP	Diag-nosis	SPECT	Neuro-Psych	PM
DAT	X	X	X	O
Controls	O	O	O	O

Table 2.1 Frequency of Data Acquisition. X - six monthly, O - once only.

The frequency with which each of these datasets is collected differs in some of the study groups, as shown

in Table 2.1. While six-monthly review data is generated from the DAT group, the normal group is assessed once only. The frequency of data collection in other dementia groups noted above follow that of DAT.

2.4 ALZHEIMER DATABASE

Each of the different datasets generated in the project are collected together and stored in a central database on the University of Glasgow mainframe ICL 3980 series computer. The Scientific Information Retrieval (SIR) database package was used to construct the Alzheimer database. SIR follows the relational model for data structures, with different datasets relating to the same entity (e.g. subject), stored in a number of separate tables and cross referenced across tables using a key identifier. A unique case identification number is used for this purpose in the Alzheimer Project, to relate the various data collected on any given subject across tables.

Study data is collected at three locations in Glasgow including Gartnavel Royal Hospital (GRH), Southern General Hospital (SGH) and the Wellcome Surgical Institute (WSI). This is illustrated in figure 2.2. At each location the data is entered manually onto record forms which are sent to the database manager. This is checked, and queried if necessary, prior to entry into computer files. When the data is 'clean' it is loaded onto the Alzheimer project database.

For the serial reassessments in the DAT group a second key identifier is used. The date on which the data was collected is used to distinguish these data. Contiguous data from different tables are matched according to these dates. For this, a time window is necessary, as data in different tables are not always collected on the same date.

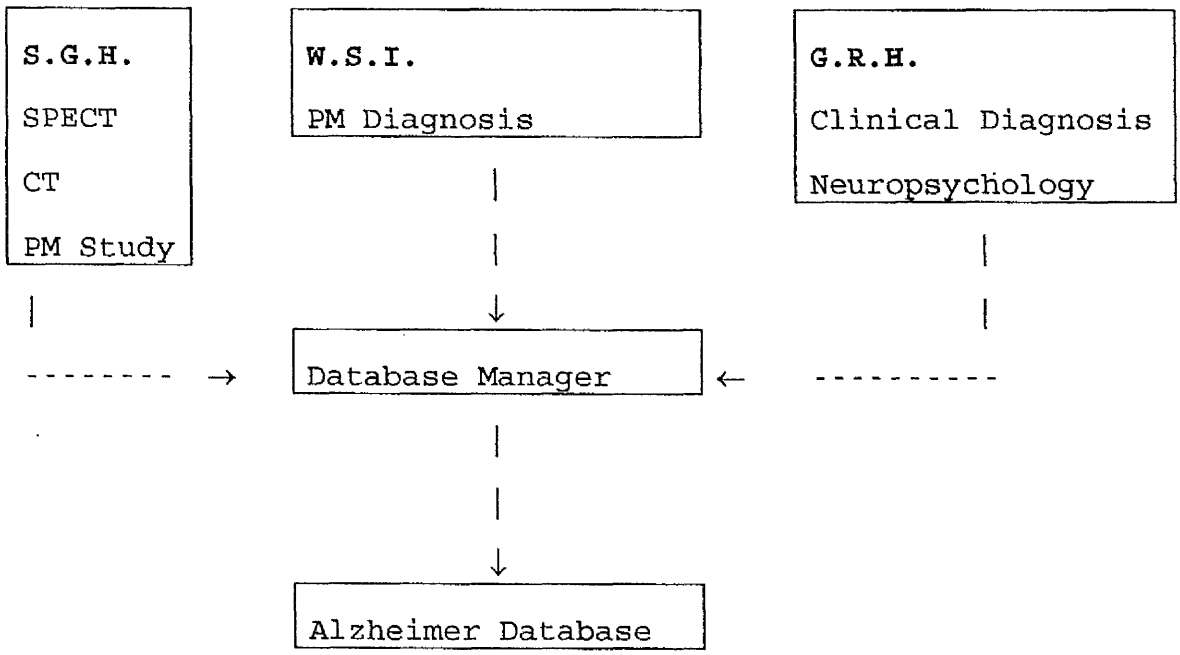


Figure 2.2 Data Flow through Project.

2.5 FUNCTIONAL SPECT IMAGING

Imaging is carried out on a dedicated Novo 810 SPECT imager. A circular gantry of gamma cameras are placed around the subject's head. During the imaging session

these cameras rotate so as to give full coverage of photon emission produced by the introduction of the radioactive isotope into the subject's blood stream. The radioactive isotope used in the study is a technetium based compound known as Ceretec or ^{99M}Tc-HMPAO.

The ceretec is administered intravenously under resting conditions 5-60 minutes prior to imaging with the imaging session usually lasting around 20-25 minutes. During this time five transverse slices (tomograms), 12 mm thick, are obtained in a plane parallel to the orbital-meatal (OM) line and at positions approximately +30, +40, +50, +60 and +70mm superior to this line. Figure 2.3 shows an illustration of two such slices. Each slice is scanned for three minutes. Two of the five images, obtained from the scanning, are then used for the purpose of ROI data extraction as outlined in section 1.3. These are chosen on the basis of best alignment through a number of anatomical structures identified from a standardised stereotactic brain atlas. The lower of the images, referred to as the `standard' slice is approximately +40mm above the OM line (Figure 2.3). For the record this is defined by the presence of the putamen, thalamus and occipital cortex and by the absence of cerebellum. The second image, referred to as the `upper' slice, is usually +70mm above the OM line (figure 2.3) and is identified as immediately superior to the corpus callosum.

Regions of interest are outlined manually on these images by an imaging technician, using a light-pen

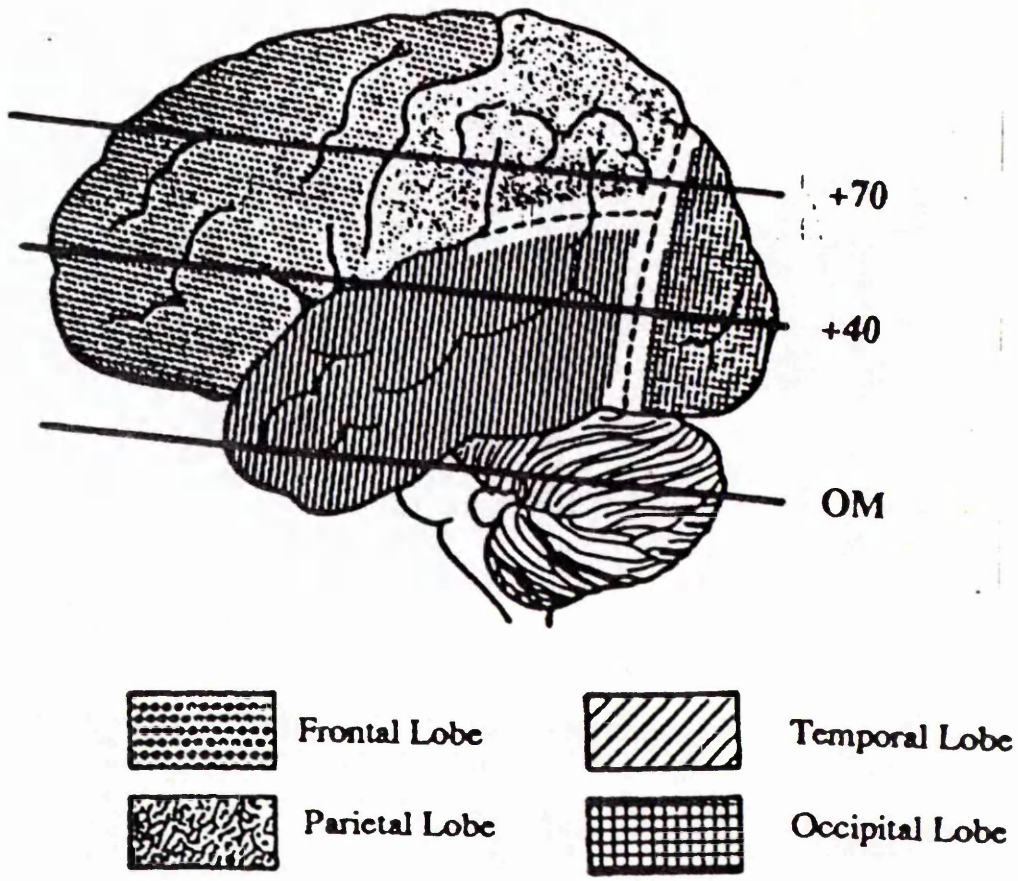
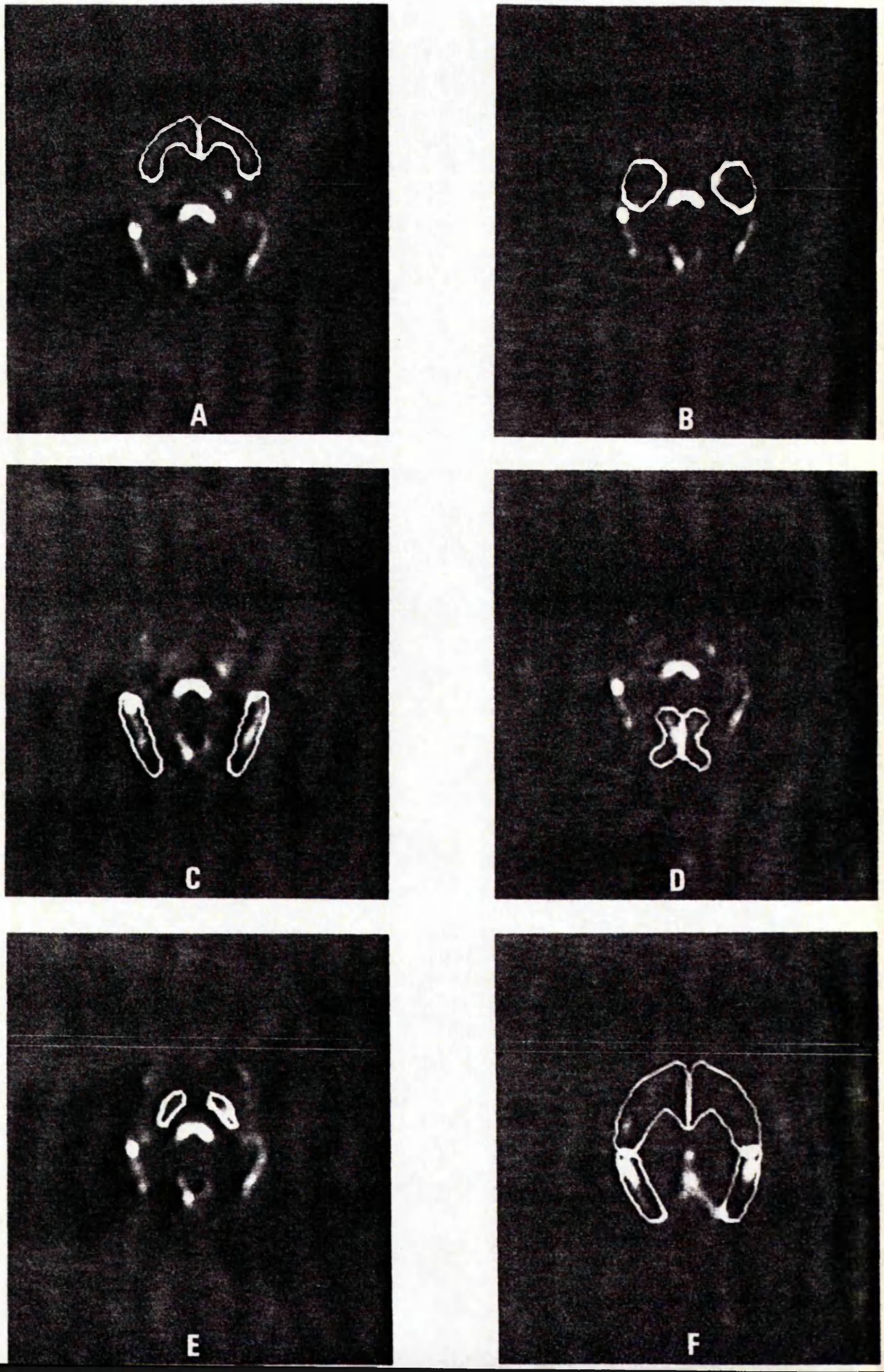


Figure 2.3 Side view diagram of the brain.

Figure 2.4 Images with ROI used in Alzheimer study superimposed. Regions in 'standard' slice are A. Frontal, B. Temporal, C. Posterior-Temporal, D. Occipital, E. Bassal-ganglia. Regions in 'upper' slice are F. Higher-Frontal & Parietal.


A
B
C
D
E
F

REGION NO.	REGION NAME	ABBREVIATION
1	Right Frontal	RF
2	Left Frontal	LF
3	Right Higher-Frontal	RHF
4	Left Higher-Frontal	LHF
5	Right Temporal	RT
6	Left Temporal	LT
7	Right Parietal	RP
8	Left Parietal	LP
9	Right Posterior-Temporal	RPT
10	Left Posterior-Temporal	LPT
11	Right Basal-Ganglia	RBG
12	Left Basal-Ganglia	LBG
13	Right Occipital	LO
14	Left Occipital	RO

Table 2.2 Region of Interest Numbers, Names and Abbreviations for reference in statistical analysis.

referencing the same stereotactic atlas mentioned previously. Figures 2.4 a-f show the approximate locations of the regions studied in the Alzheimer project. A measure of the activity is extracted from these regions. This is the mean counts/pixel. In addition, an outline of the whole image is taken.

The names of the regions and abbreviations (for reference in subsequent analysis) are given in Table 2.2.

2.6 NEUROPSYCHOLOGICAL STUDY OF CEREBRAL FUNCTION

Neuropsychology is the science concerned with the study of the organisation of cerebral function in the living human brain. As such, it is concerned with the nature of mental processing and it's relationship to, for example, cognition and behaviour. In order to study aspects of dysfunction the subject sits a number of simple neuropsychological tests. The test illustrated in Figure 2.5 is used to assess language deficits resulting from the disease process. As can be seen this is an easy test for someone with little or no language problems. As a measure of the difficulties with language, the number of errors is used for evaluation and analysis. A further example is a test used to assess memory deficits. This involves the immediate recollection of a list of items having only seen them previously for a few seconds. Many such tests are used in the Alzheimer Project.

GRADED NAMING TEST (McKENNA & WARRINGTON)

Time: 30 seconds per item
Discontinue after 5 consecutive failures.

WORD	LATENCY	ERROR	DETAILS(vis./rel/other.)	SCORE
BED				
CLOCK				
TAP				
KEY				
PIANO				
CIGARETTE				
KANGAROO				
SCARECROW				
BUOY				
THIMBLE				
HANDCUFFS				
TWEEZERS				
CORKSCREW				
SPORRAN				
TASSEL				
SUNDIAL				
CHOPSTICKS				
PERISCOPE				
BOAR				
BLINKERS				
MONOCLE				
TURTLE				
TRAMPOLINE				
BELLOWS				
SHUTTLECOCK				
ANTEATER				
PAGODA				
RADIUS				
LEOTARD				
MITRE				
YASHMAK				
SEXTANT				
CENTAUR				
COWL				
TUTU				
RETORT				

TOTAL SCORE	[ ][ ]	TOTSCOR 60-61
TOTAL TIME (SECS)	[ ][ ][ ]	TOTTIME 62-64
ERRORS - Visual	[ ][ ]	VISERR 65-66
Relational	[ ][ ]	RELERR 67-68
Total	[ ][ ]	TOTERR 69-70

Figure 2.5 Neuropsychological test sheet for Graded Naming Test.

2.7 POST-MORTEM STUDY OF BRAIN TISSUE

As mentioned previously, neuritic plaques and neurofibrillary tangles are widely regarded as the main pathological markers of DAT. For this reason consent is routinely sought from relatives, to carry out post-mortem study of brain tissue. If permission is given, then at death the brain is removed and alternate slices stored at the Wellcome Surgical Institute and Southern General Hospital. These are stored for some time in order to eliminate any danger of contamination from contagious diseases, such as Kreutzfeld-Jacob syndrome, which can be contracted from handling brain tissue. Eventually, samples of tissue are removed for microscopic study and, from these quantitative estimates of the plaque density obtained.

BRAIN ROI	1	2	3	.....
BLOCK	1	1	1	.....
SECTION	1	1	1	.....
REFERENCE POINT	1 - 6	1 - 6	1 - 6	.....

FIGURE 2.6 Sampling frame for plaque quantitation

In the study a hierarchical sampling procedure is used to quantify plaque density in different areas of brain tissue. A brief summary of the procedure is given with the aid of the sampling frame in Figure 2.6. Coronal slices (1cm thick) of brain tissue are cut. Blocks are taken from these corresponding to the particular anatomical area of interest. A microscopic section is sliced from this and stained so as to highlight plaques from background features. A NOVO 810 image analyser is linked to the microscope and used to place six small reference fields ($0.645mm^2$), according to a specified protocol, in different areas of the slide. Within these fields the number of plaques is counted and the average used as an estimate of the density in the slice and hence the block and finally the ROI. Quantitative estimates of plaque density are obtained from a number of ROI in this way. Some of these regions correspond with those used in SPECT.

Confirmatory diagnosis of DAT can then be made using this data. For this, a common approach among neuropathologists (Wieszniewsky, 1988), is to use the Khatchaturian criteria (Khatchaturian, 1985).

CHAPTER 3

STATISTICAL IMAGE ANALYSIS: ISSUES AND PROBLEMS OF INTEREST WITH APPLICATIONS TO A SPECT DATASET

Questions of interest posed of imaging datasets are broadly classified into problems of within group and between group analysis of patterns of regional activity. Typical within group problems include investigations of regional means for evidence of hemispheric asymmetry or more complex hemisphere x region interactions. At a biological level other analysis is involved in studies of inter-regional association. Between group problems are concerned with the comparison of mean profiles and correlation structures as well as problems of discrimination.

A large proportion of the work found in neuroscientific literature on these problems involves the use of univariate statistical methods. In this chapter we will illustrate some of the difficulties of interpretation that can occur with the univariate approach.

3.1 A SPECT DATASET AND SOME NOTATION

Mean counts/pixel data were recorded in 14 regions in a group 29 control subjects and 79 cases with a clinical diagnosis of dementia of the Alzheimer type. A description of the method of collecting these data and of ROI extraction is given in section 2.5. A sample of these

data is given for four subjects in Table 3.1 below.

In the analyses that follow we will assume that these data arise from an unknown multivariate normal probability distribution with group mean vector $\underline{\mu}_j = E(\underline{X}_{ij})$ and covariance matrix Σ_j, for the ith subject $i=1,\ldots,n_j$ in $j=1,..J$ groups . In ensuing sections, questions of interest will focus on aspects of $\underline{\mu}_j$ and Σ_j. Distributional assumptions will be assessed at various stages.

SUBJECT	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	150	159	139	143	158	156	166	161	166	162	128	131	168	177
2	206	210	211	215	225	227	229	229	245	245	268	273	245	244
3	173	180	176	176	197	201	212	211	205	205	154	170	198	193
4	82	82	92	89	89	97	104	111	98	99	84	82	109	102

Table 3.1 Listing of four subjects from the SPECT dataset.

3.2 SOME BACKGROUND TO SPECT DATA

A feature of quantitative ROI data is the presence of large 'random' scalar differences in data vectors between subjects and even between scans on the same subject. This can be seen to some extent in the four data

vectors shown in Table 3.1. There are indications of large differences in the magnitude of the counts across the individuals. On plotting the individual profiles (Figure 3.1) the global difference is quite clear.

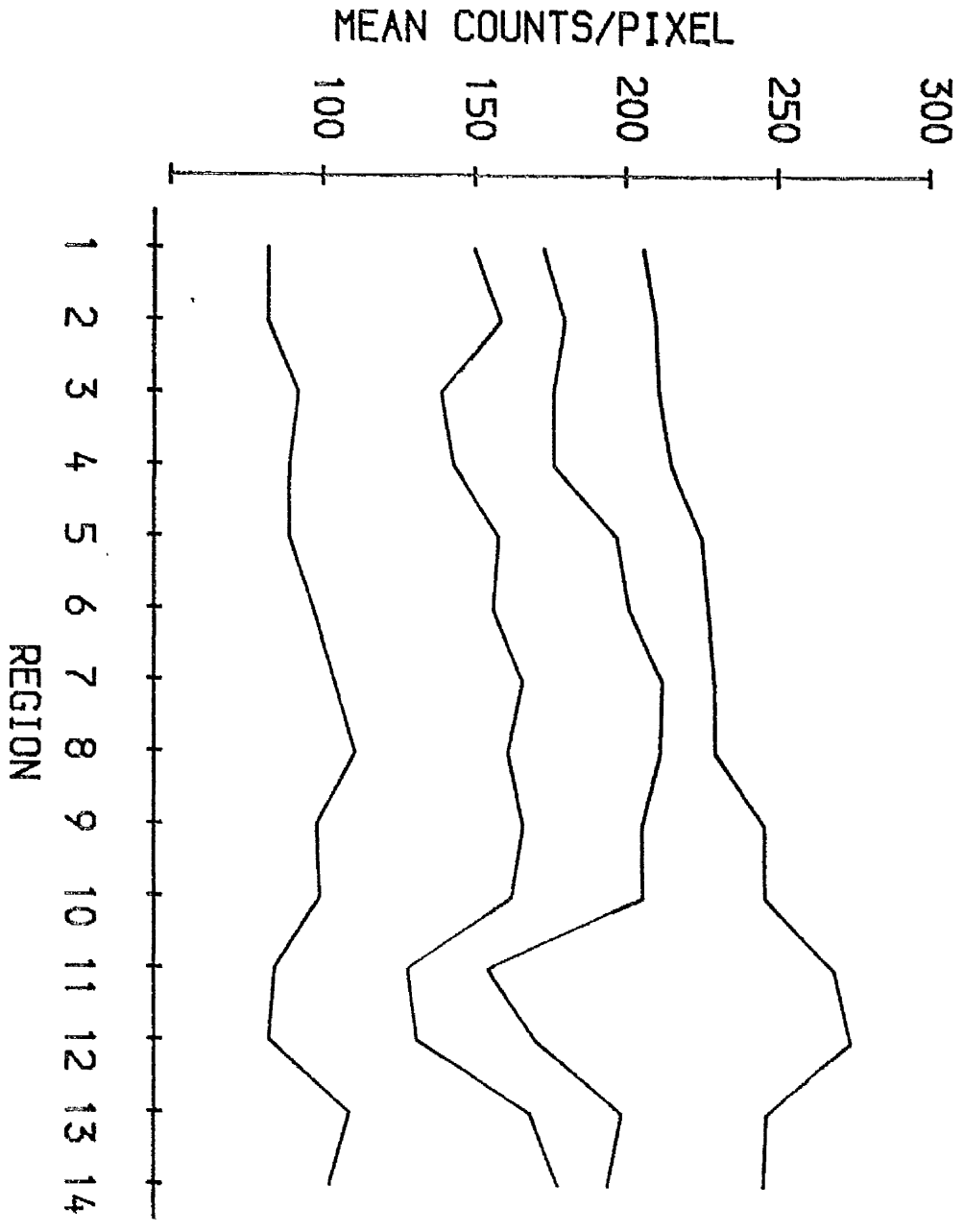
There are a number of reasons for the presence of the scalar effect in imaging datasets. The external and/or internal stimulation of the subject during an imaging session will determine, to some extent, the level of activity in the subject's brain. Since the level of activity is linked to cerebral demand (Kuschinsky, 1981), this will have an impact on the amount of tracer transported to the brain. Any anxiety experienced by the subject during an imaging session can cause this stimulation. The increased demand may result in greater number of reconstructed cell counts. At the other extreme, if too little tracer is taken up into the organ of interest, few emissions may be detected. This may in turn lead to sparse data sets available for reconstruction.

3.3 NORMALISATION

In the analysis of these SPECT data there are two basic approaches taken to account for the scalar effect.

One is to factor in the scalar effect into a model of the data. Regression analysis where the subject mean, or some other global measure, is used as a covariate has been used. Another is to carry out a repeated measures

Figure 3.1 SPECT profiles for four normal subjects.



type analysis (see chapters 4). The other more common approach is somewhat cruder and simpler. In this approach the scalar effect is explicitly removed from the data prior to any statistical analysis.

This latter approach is referred to, in the literature, as normalisation. It involves a simple transformation of each ROI value into ratios or differences from some other value of the subject's vector. The two common forms of normalisation are as follows. The first is to normalise the ROI data to another regional value - putatively 'spared' from the disease process under study. For the data vector $\underline{X}_i$, for subject i, normalisation to region X_r, results in a new (p-1) vector data vector $\underline{Y}_i$ of normalised regional activity of the form:

$$\underline{Y}_i = (X_1/X_r,\ X_2/X_r,\ \ldots\ldots,\ X_{r-1}/X_r)$$

The other form of normalisation is to a global measure of activity (e.g. subject mean) or other global measure (e.g. the slice average). A new (p) vector $\underline{Y}_i$, of normalised activity is obtained of the form:

$$\underline{Y}_i = (X_1/X.,\ X_2/X.,\ \ldots\ldots\ldots\ldots,\ X_p/X.).$$

3.4 HEMISPHERIC ASYMMETRY

Functional symmetry in the human brain is aspect of interest to investigators studying neurological

dysfunction (Engel, 1982; Rhodes, 1983; Perlmutter, 1985, 1987). For regions, with both right and left measurements, the traditional approach to comparison of hemispheric mean levels of activity is to transform to bilateral ratios (or differences or percentage differences) and then to carry out a univariate (region by region) one-sample analysis (Perlmutter, 1985, 1987). Note there is no need to pre-normalise the data for this analysis, because the transformation has the effect of adjusting for the scalar effect between subjects.

This analysis is illustrated below for each of our study groups separately. Right/left ratios were calculated for each region. Means, standard deviations and 95% confidence intervals for these ratios were calculated (Table 3.2a,b). A test of the null hypothesis that regions are symmetric was carried out using a one sample t-test.

From table 3.2a we can detect evidence of asymmetry in four regions in the DAT group; there being greater activity in the right side than the left side on average. In contrast there is little evidence of any difference, apart from the occipital region, in the normal group.

3.5 FUNCTIONAL ASSOCIATION AMONG REGIONS

When studying the pattern of functional associations between regions of the brain many investigators have used correlation analysis (Clark et al, 1984; Horwitz et al, 1984; Metter et al, 1984a,b; Bartlett et al, 1988).

REGION	MEAN	STD.DEV.	95% C.I.	P-VALUE
Frontal	1.018	.055	(1.006 , 1.031)	.0040
Hi-Frontal	1.006	.065	(0.992 , 1.021)	.38
Temporal	1.019	.070	(1.003 , 1.034)	.020
Parietal	1.023	.086	(1.004 , 1.042)	.020
Post-Temp	1.011	.084	(0.992 , 1.029)	.26
Basal-Gang	0.994	.105	(0.970 , 1.017)	.61
Occipital	1.021	.068	(1.003 , 1.036)	.0063

(a) DAT Group

REGION	MEAN	STD.DEV.	95% C.I.	P-VALUE
Frontal	1.010	.040	(0.995 , 1.025)	.20
Hi-Frontal	1.002	.033	(0.992 , 1.021)	.79
Temporal	1.007	.057	(0.986 , 1.029)	.50
Parietal	1.003	.066	(0.978 , 1.028)	.79
Post-Temp	1.010	.068	(0.984 , 1.035)	.46
Basal-Gang	0.990	.087	(0.957 , 1.023)	.56
Occipital	1.012	.031	(1.000 , 1.023)	.053

(a) Normal Group

TABLE 3.2 Summary data and p-values from hypothesis H: $\mu_r/\mu_l=1$ within (a) DAT group and (b) normal group.

HF	.99					
Temp	.98	.99				
Par	.98	.97	.98			
PT	.96	.96	.94	.96		
BG	.95	.95	.94	.96	.98	
Occ	.97	.96	.97	.99	.96	.97
	Fr	HF	Temp	Par	PT	BG

TABLE 3.3 Sample correlation matrix of hemispheric region sums.

The sample correlation matrix between regions in the normal group is given in Table 3.3. To simplify the presentation, only correlations between regions where the data was transformed to hemispheric sum counts (i.e. RHS+LHS) are presented. The correlation coefficients are large and positive, which would suggest very strong association between regions. However, we know already from section 3.2 that the raw data is characterised by large random variation between subject data vectors, which as we will see in chapter five, is really the dominant feature of these correlations.

As with the analysis of means, a common solution is to normalise the data. Correlation matrices for various forms of normalisation are presented in Table 3.4(a)-(c). These include (a) normalisation to a subject mean, (b) to the occipital region and (c) partial correlation approach

(a)

HF	.66					
Temp	.29	.49				
Par	-.42	-.47	.22			
PT	-.20	-.33	-.77	-.33		
BG	-.65	-.53	-.68	-.16	.39	
Occ	-.44	-.60	-.05	.43	-.30	.15
	Fr	HF	Temp	Par	PT	BG

(b)

HF	.87				
Temp	.67	.79			
Par	.32	.37	.53		
PT	.46	.44	.08	.27	
BG	.08	.20	-.11	.09	.62
	Fr	HF	Temp	Par	PT

(c)

HF	.75					
Temp	.18	.38				
Par	-.41	-.39	.24			
PT	-.40	-.45	-.61	-.15		
BG	-.50	-.53	-.61	-.35	.41	
Occ	-.49	-.57	-.17	.30	-.15	.14
	Fr	HF	Temp	Par	PT	BG

TABLE 3.4 Correlation Matrices for data (a) normalised to subject mean, (b) normalised to occipital activity and (c) partial correlation coefficients.

using the subject mean as a covariate. The third of these (table 3.4 (c)) illustrates an alternative form of adjustment which has been used frequently to study interregional correlations. This uses the subject mean as as a covariate for calculating partial correlation coefficients (Horwitz 1984, 1987, 1990) between regions.

The differences between the correlation matrix in table 3.3 and those in tables 3.4 a-c are quite dramatic. Equally, the differences between some of the correlation matrices in table 3.4 are substantial. In particular 3.4 (a) and (c) are very different in appearance to (b) with 3.4 (a) & (c) bring quite similar.

While each of the analses above are valid, it is clear that substantially different impressions of the relationship between regions can be drawn from the data using different approaches to remove the subject effect.

3.6 BETWEEN GROUP COMPARISONS

The comparison of regional mean profiles between groups figures very prominently in research papers, with the univariate approach being the most common approach for the analysis. This necessarily involves the removal scalar effect by normalisaton (Risberg, 1985; Burns et al, 1989; Montaldi et al, 1990) may differ from study to study, even if investigating the same disease groups. A comparison regional means between our normal and DAT groups was carried out using different forms of normalisation (Table 3.5).

REGION	(1)	(2)	(3)	(4)
Frontal	.44	.029	.021	.26
H-Frontal	.53	.04	.035	.14
Temporal	.37	.013	.0045	.68
Parietal	.28	.015	.0000	.66
Post-Temp	.40	.044	.0001	.36
Basal-Ganglia	.70	-	.81	.0011
Occipital	.76	.92	-	.0000

TABLE 3.5 P-values from two-sided 2-sample t-tests. The '-' indicates that comparison does not apply. Columns are (1) Raw Counts data, (2) normalised to basal-ganglia (3) normalised to occipital and (4) normalised to subject mean.

Analysis is carried out on raw mean counts/pixel data (1), after normalisation to occipital activity (2), basal-ganglia activity (3) & finally to a global subject mean (4). As in Section 3.5, regions with bilateral data were first transformed to hemispheric sums. For each region, the null hypothesis of zero mean difference between groups was tested using 2 sample t-tests. The results are summarised with p-values in Table 3.5.

There is no evidence of regional difference based on raw data (column 1). After normalisation a number of

differences between groups are shown, especially in columns 2 & 3 and less so in column 4.

The results in column 1 are of course due to the large random variation contributing to large region variances. In Alzheimer's disease the occipital and perhaps basal-ganglia regions are putatively 'spared'. Thus the conclusions from analysis of data normalised to these regions give similar conclusions when each comparison is made at the nominal 5% level of significance. However, given the large number of comparisons made however, some form of adjustment should generally be considered, to control overall error rates. After Bonferroni adjustment to the analysis in each column (e.g. with individual significance assessed at the nominal 0.05/6% level in columns 2 & 3), many of the comparisons are no longer statistically significant. Note especially the effect on results in the basal-ganglia normalised column. After normalisation, none of the regions are considered significant. For column 4, there is an understandable tendency to interpret the group differences as a property of the two significant regions. However, this would be a mistaken conclusion. Having noted these two regions to be spared, a subject mean in the disease group will be relatively speaking, lower (by comparison) than in the control group. Hence division by the subject mean in the DAT group will increase the regional ratios, causing big differences between group means resulting in the pattern of significant differences seen here.

3.7 FURTHER CONSIDERATIONS ABOUT ROI DATA

An aspect in imaging data which is worth considering before moving on to more complex approaches to the analysis is the scale of measurement or metric used in the analysis. Clark et al (1985) and Moeller et al (1987) noted that the within-subject standard deviation among regional measurements on a given subject increased linearly with the within-subject regional mean. If regions were randomly related then this would clearly indicate the need for some transformation of the raw count data. However, since regions are structurally related, it is not clear what this actually tells us. In our dataset the correlation coefficients between subject means and standard deviations is 0.62 in the normal group and 0.47 in the DAT group.

3.8 DISCUSSION

The analyses presented in this chapter are not dissimilar to those found in much of the experimental neuroimaging literature. In both the between group comparisons of means and in the correlation analysis of functional associations, we saw how easily different biological conclusions could be drawn from the same dataset. In addition, although these analyses used univariate statistical methods, the nature of the subject matter warranted a multivariate interpretation, which was

attempted in terms of ratios of random variables. In problems involving several disease groups the normalisation approach may be even more complicated.

In our dataset the number of regions is, by the standards of the studies described in table 1.1 , very small, while the number of subjects in each group (especially the DAT group) is comparatively large. In studies with many regions the univariate analysis of means will involve many comparisons. Consequently, adjustments to comparison error rates will be required. In this respect Bonferroni adjustment is common. This can result in conservative joint significance levels, and make meaningful differences between groups and regions difficult to detect. Alternative adjustments using the Sidak approach may be more appropriate (MacCormack, 1991)

Finally, it is worth noting some similarities between ROI data and other data types. Pixelated individual images have immediate parallels with datasets obtained in spatial problems (Diggle, 1983). The complication of the scalar effect in ROI data is analagous to the size factor in allometric data sets (Mardia, Kent & Bibby, 1979). In later chapters we will also see links with problems encountered in compositional data analysis (Aitchison, 1986) through the use of normalisation.

CHAPTER 4

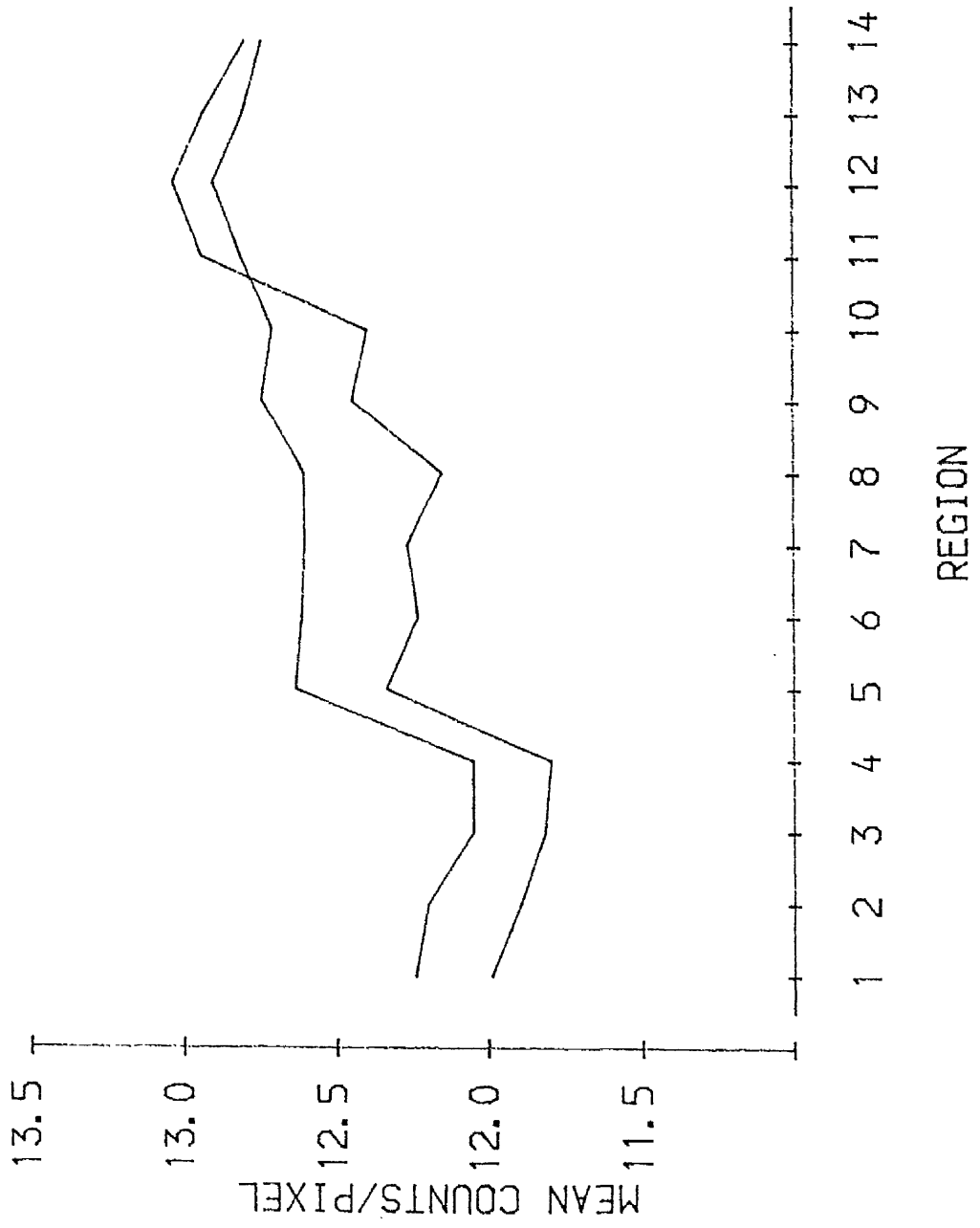
ANOVA APPROACH TO ROI PATTERN ANALYSIS

While the univariate approach to the between group problem provides useful information about group differences, the results can be difficult to interpret. In addition, this approach can involve a large number of comparisons which, after correcting for the family experimentwise error rate, can make significant differences very difficult to detect. Even with the relatively small number of regions in our study, the effect of Bonferroni adjusted significance levels had a big impact in the analysis of basal-ganglia normalised data. A further limitation of the univariate approach is that it took no account of the inter-regional correlation structure.

Let us reconsider this same data now. Regional unadjusted mean profiles of square root data are plotted for the two groups in figure 4.1. Justification for this transformation will be given in a later section. The two profiles are approximately parallel from regions 1 to 10 with a crossover of profiles from regions 11 to 14. We are essentially interested here in testing the hypothesis $H_0 : \mu_1 = \mu_2$, against the alternative $H_1: \mu_1 \neq \mu_2$ and, where applicable, follow-up analysis comparisons of regions.

Because of the nature of the data repeated measures ANOVA techniques have been used to study imaging datasets (Kelly et al, 1982; Tyler et al, 1988; Worsley et al,

Figure 4.1 Group mean profiles based on square root counts.



1991). We now describe two forms of the repeated measures ANOVA approach; the univariate approach and the multivariate approach.

4.1 UNIVARIATE REPEATED MEASURES ANOVA

Univariate repeated measures ANOVA techniques (Crowder & Hand, 1990) are commonly used to analyse observations which are measured in sequence in time or in space on the same experimental or observational unit e.g. subject or animal. The repeated measures ANOVA is essentially a split-plot analysis where time replaces treatment as the within main-plot (here subject) factor of interest. The basic difference between the two is that time is a fixed effect factor as opposed to a random factor in the traditional split-plot paradigm. It is probably no surprise then to see this technique used to investigate regional mean patterns in neuroimaging data in PET (Tyler et al, 1988; Worsley et al, 1991) and Autoradiographic studies (Kelly et al, 1982).

We assume a model of the form

$$X_{ijk} = \mu_{jk} + P_{i(j)} + \varepsilon_{ijk}$$

where μ_{jk} represents the average response in the k^{th} region (k=1,..,p) of group j, $P_{i(j)}$ denotes the random effect due to the i^{th} subject in group j and ε_{ijk} represents residual error in the k^{th} region. In addition it is assumed that the $P_{i(j)} \sim N(0,\sigma^2)$ and $\varepsilon_{ijk} \sim N(0,\tau^2)$.

The μ_{jk} can be further decomposed into for example region, group and region x group effects. The region effect is included to describe any differences in the mean levels of activity among regions while the region x group interaction represents any differences in the regional mean pattern across groups. We are primarily interested in the group x region interaction.

SOURCE	D.F.	F	Unad-P	GG	HF
Group	1	0.28	0.5997		
B-Subject Error	106				
Region	13	39.13	0.0000	0.0000	0.0000
Region x Group	13	3.44	0.0000	0.0048	0.0039
W-subject Error	1378				

TABLE 4.1 : Repeated measures ANOVA analysis of profiles. Adjusted p-values are calculated using Greenhouse-Geisser (GG) and Huyn-Feldt (HF) methods.

A repeated measures analysis was carried out for the data in section 3.1, using the BMDP program P2V, including one grouping factor (Group) and one level factor (Region). The results are summarised in Table 4.1 above.

The overall group comparison uses between subject error which, because of the large subject effect, is not significant. On the other hand, the within subject part of the analysis is based on contrasts of the form $Y_{ijk}=X_{ijk}-X_{ij.}$. By subtracting the subject mean, the within-subject analysis is therefore analogous to analysis of normalised data.

Restricting interpretation to the column headed 'Unad-P', the following conclusions can be drawn. The regional effect is very strong indicating statistically different levels of activity across regions. The group x region interaction test is also highly significant indicating differences between groups in the levels of activity in some of the regions. In effect, group mean profiles are not parallel.

A feature of this approach is that the validity of these univariate F-tests depends on restrictive assumptions about the form of the covariance structure of repeated measurements. For these unadjusted within-subject tests to be valid, it is necessary that group covariance matrices are similar $\Sigma_1=\Sigma_2(=\Sigma)$ and that Σ satisfies the sphericity assumption. Here we assume that the variance of all pairwise differences between variables $V(X_i-X_j)$, for $(1\leq i<j\leq p)$, are equal. A special case of this is known as compound symmetry. This is when the variances $V(X_i)$, $(i=1,..p)$, of all variables are equal and the covariances $C(X_i,X_j)$ are equal. In this instance the repeated measurements will be equicorrelated. In these data the test of sphericity,

based on the pooled sample covariance matrix, is strongly rejected (p=0.0000). Thus, the unadjusted F-tests are not strictly valid. However, this does not mean that the univariate analysis itself is invalid.

Adjusted univariate F-tests can still be made, using correction factors developed by Greenhouse-Geisser (1959) and Huyn-Feldt (1976). These correction factors involve a reduction to the numbers of degrees of freedom used to look up the tabulated F-values. This increases the F-value for comparison with the F-statistic and results in an increase to the observed p-value. These adjusted p-values are included in standard output for this type of analysis. In the analysis above the adjusted p-values are also given, in columns headed 'GG' and 'HH' of table 4.1. Even after adjustment, the p-values are much lower than the 5% levels. Hence, our conclusions remain unaltered.

Before moving on, it is worthwhile pausing to consider the relative size of the error terms generated from this analysis (not given in Table 4.1). The ratio of random between-subject error variation to total random error is 1.12614/(1.12614 + 0.00735)=0.993. Thus between subject variation accounts for 99.3% of the total variability in this data.

4.2 MULTIVARIATE REPEATED MEASURES ANOVA

An alternative approach is to view the regional measurements for any subject as a multivariate vector and to use multivariate analysis of variance (MANOVA)

techniques (Crowder and Hand, 1990). This is the multivariate repeated measures analysis of variance.

The multivariate model is written as

$$\underline{X}_{ij} = \underline{\mu}_j + \underline{\varepsilon}_{ij}$$

where $\underline{\mu}_j$ denotes the p-vector of regional activity in the jth group and $\underline{\varepsilon}_{ij}$ denotes the vector of residual error. In addition it is assumed that the vector of errors $\underline{\varepsilon}_{ij}=(\varepsilon_{ij1},\ldots,\varepsilon_{ijp})$ are $MVN(0,\Sigma)$.

Hypotheses, under this model, are conveniently written in the form

$$H: C\underline{\mu}M=0$$

where C represents the ((g-1) x g) matrix of group contrasts, $\underline{\mu}$ is the p-vector of regional means and M represents the (p x (p-1)) matrix of contrasts among regions. With suitable choice of M, a wide variety of hypotheses can be addressed. The simplest case is when M=I - the identity matrix. In the event of a significant difference further comparisons can be used to assess the parallelism of profiles. For this we need to choose a form of difference matrix M, that compares regional contrast(s) among groups. There are many difference matrices that could be chosen here. A contrast of all variables with a single region can be expressed in matrix form as $M_1 = [\ I_{p-1} : -j_{p-1}\ ]$, which is the (p x (p-1)) identity matrix augmented by a (p x 1) column vector of -

1's. In the two group case this simplifies to H: $M^T(\mu_1-\mu_2)=0$ which can be tested using Hotelling's T^2 test.

The equality of mean profiles in figure 4.1 was tested using this approach. The results are given in Table 4.2. We can conclude from these results that there is insufficient evidence to reject the null hypothesis. Using the averaged right and left occipital activity, regional data was differenced and groups compared once again (Table 4.2). Again there is no evidence of differences between groups.

This is surprising in view of the highly significant interaction effect in the RM ANOVA analysis. The reason for this is unclear at this stage. The sample sizes are very different and the number of dimensions is high. However, this will only be a problem if covariance matrices are unequal. We will come back to this later.

TEST	D.F.	T^2	P-VALUE
Equality	14,93	20.36	0.2375
Parallelism	12,95	10.03	0.6998

TABLE 4.2 MANOVA comparison of mean profiles.

There is very little evidence of the use of MANOVA

in the analysis of ROI data - understandable given the descriptions of the dimensions of typical study datasets given in chapter 1. A rare example can be found in Prohovnik (1988), dealing with the analysis of Xenon washout data .

4.3 HEMISPHERIC ASYMMETRY

In instances where measurements are available in both hemispheres for some or all brain regions it makes sense to incorporate this structural information into the analysis.

In a MANOVA analysis this can be achieved by transforming to hemispheric sums and differences and carrying out two separate analysis. For our data this involves comparisons of two seven dimensional mean vectors between groups - instead of a single 14 dimensional comparison. The results of these tests for equality and parallelism are given in Table 4.3. While hemispheric sums are significantly different the hemispheric differences are not significant. The follow-up test of parallelism on the hemispheric sums was carried out using differences to occipital activity as in section 4.2. This is also significant giving evidence of a group x region interaction. This is surprising in view of the comparison of the full mean vectors.

TEST	D.F.	T^2	P-VALUE
Equality			
Sums	7,100	18.32	0.02
Differences	7,100	2.74	0.92
Parallelism			
Sums	6,101	18.31	0.01

TABLE 4.4 MANOVA comparison of mean profiles.

In the univariate RM model, the hemisphere effect can be easily incorporated in to the analysis by including a within-subject factor for hemisphere. This produces an model with a within-subject factorial structure. This is illustrated for the data above. A full univariate analysis was carried out on our dataset with the within part of the analysis given in Table 4.4.

The region x group interaction is still very strong here, both with and without the adjustment. There is also some evidence of a region x hemisphere interaction although this is somewhat equivocal in these data, since the p-values are close to the nominal 5% level. Indeed the Greenhouse-Geisser correction takes the p-value to just over the 5% level. This illustrates a difficulty

SOURCE	D.F.	F	Unad-P	GG	HF
Hemisphere	1	1.55	0.2162		
Hemisphere x Group	1	0.86	0.3563		
W-subject Error (1)	106				
Region	6	47.72	0.0000	0.0000	0.0000
Region x Group	6	4.13	0.0004	0.0046	0.0039
W-subject Error (2)	636				
Region x Hemisphere	6	2.34	0.0304	0.0502	0.0467
Region x Hemisphere x Group	6	0.29	0.9441	0.8982	0.9066
W-subject Error (3)	636				

TABLE 4.4 : Repeated measures ANOVA analysis of regional means. Adjusted p-values are calculated using Greenhouse-Geisser (GG) and Huyn-Feldt (HF) methods.

that can arise when the sphericity assumption is violated. The difficulty can be in choosing an appropriate p-value. The Huyn-Feldt adjustment is generally believed to be less conservative than the Greenhouse-Geisser approach. In this instance, the Greenhouse-Geisser p-value is so close to the nominal level as to warrant a common sense interpretation as a significant effect.

4.4 FOLLOW-UP ANALYSIS OF PROFILES

Following a significant overall test result e.g. region effect or group x region interaction, a follow-up analysis will often be desired, to determine the nature of differences between profiles. It clearly makes little sense to compare regions between groups individually, as this would involve between subject variation (Table 3.2, column 1). Alternative approaches are required, some of which are described below.

4.4.1 Multiple Comparisons

In the event of a significant group x region interaction, follow-up comparison of profiles might be carried out using multiple comparison procedures. Since it is pointless to compare regions individually, we need to look at contrasts among regions for comparison between groups.

Let $Y_{rs}=(\mu_{r1}-\mu_{s1})-(\mu_{r2}-\mu_{s2})$ denote the contrast between the rth and sth regional means in group 1 and group 2. This can be estimated using the sample regional means. Assuming that Σ satisfies the sphericity assumption an unbiased estimate for the variance $V(Y_{rs})$ would involve the region x group x subject mean square error term from the RM ANOVA table. Simultaneous confidence intervals for the p(p-1)/2 contrasts could then be constructed by $\hat{Y}_{rs} \pm CV \times [\hat{V}(\hat{Y}_{rs})]^{1/2}$, where CV is an appropriate critical value.

						Variances
	0.040	0.085	0.118	0.173	0.245	0.218
0.074		0.076	0.110	0.189	0.249	0.254
0.125	0.199		0.102	0.145	0.244	0.208
0.239	0.313	0.114		0.075	0.297	0.118
0.049	0.123	0.076	0.190		0.322	0.115
0.824	0.750	0.949	1.062	0.872		0.350
0.748	0.674	0.873	0.986	0.796	0.076	
Means						

TABLE 4.7 Means and variances of contrast $Y_{rs}=(\mu_{r1}-\mu_{s1})-(\mu_{r2}-\mu_{s2})$.

However, as noted earlier, the sphericity assumption was rejected for these data. Hence the nominal confidence level for these comparisons will not be appropriate using

the above variance estimate. In this case we could use individual variance estimates $V(Y_{rs})=2(s_r^2 + s_s^2 - 2s_{rs})(1/n_1 + 1/n_2)$ for each of contrasts Y_{rs}. Assuming equal covariance matrices these would be obtained from the pooled sample covariance matrix. For our data the 7 x 6/2=21 contrasts among regional hemispheric sum means are summarised in table 4.7. $\hat{Y}_{rs}$ values are given in the lower triangle and corresponding variance estimates $\hat{V}(\hat{Y}_{rs})$ in the upper triangle.

For illustration and to assess statistical significance, a critical value was chosen from the t-distribution, with $n_1+n_2-2=106$ degrees of freedom. Because of the multiplicity of testing, Bonferroni adjusted significance levels at the $\alpha_B = .05/21$ were used. Significance was assessed in cases where the ratio $\hat{Y}_{rs}/[\hat{V}(\hat{Y}_{rs})]^{1/2}$ exceeded $t(1-\alpha_B;106)=3.1131$. None of the contrasts in Table 4.7 exceed this value. This would suggest that some other contrast(s) among regions accounts for the difference between groups observed in Table 4.4.

4.4.2 Stepwise Variable Selection in Discriminant Analysis

An alternative approach might be based on stepwise variable selection procedures in discriminant analysis. This approach is illustrated for these data using BMDP program P7M. To illustrate some important aspects about the use of this approach with these data we picked out

some steps in the selection process. Table 4.7 gives F-to-enter values for three steps in two different runs of the program. In the first run, selection was allowed to run freely with no constraints placed on selection other than to ensure all regions were forward selected before backward elimination. The initial and final F-to-enter values are given for each region in columns 1 & 2. In the second run the occipital region was forced into the model at step 1. The step 2 F-to-enter values for the remaining regions are given in column 3.

The initial values in run 1 are small because these are univariate F-tests using between subject variation. These results compare with those in column 1 in Table 3.2. Backward selection was achieved by setting F-to-enter=0. At the end of this run only the occipital and parietal regions F-to-enter values were large enough to be selected. In run 2, by forcing the occipital region into the model, we see that all regions other than the basal-ganglia now show some evidence of a difference between groups. That the parietal variable has the largest F-to-enter value means that it will be selected in the next step. Thereafter no variables would be selected - as seen in column 2. However, the F-to-enter values in column 3 are large enough to be selected, only just being smaller than for the parietal region. This

F-TO-ENTER VALUES

REGION	----- RUN 1 ------ INITIAL VALUES	FINAL VALUES	- RUN 2 - FORCED OCCIPITAL
Frontal	0.57	0.46	5.01
Hi-Frontal	0.42	3.03	3.49
Temporal	0.80	0.07	7.30
Parietal	1.17	16.43	16.43
Post-Temporal	0.69	0.52	10.97
Basal-Ganglia	0.10	3.55	0.02
Occipital	0.08	15.19	-

TABLE 4.7 F-To-Enter values from different runs of BMDP stepwise variable selection program P7M.

implies that a number of regions may actually involved in the disease process - not just occipital and parietal regions.

4.5 CHECKS ON ASSUMPTIONS

4.5.1 Normality

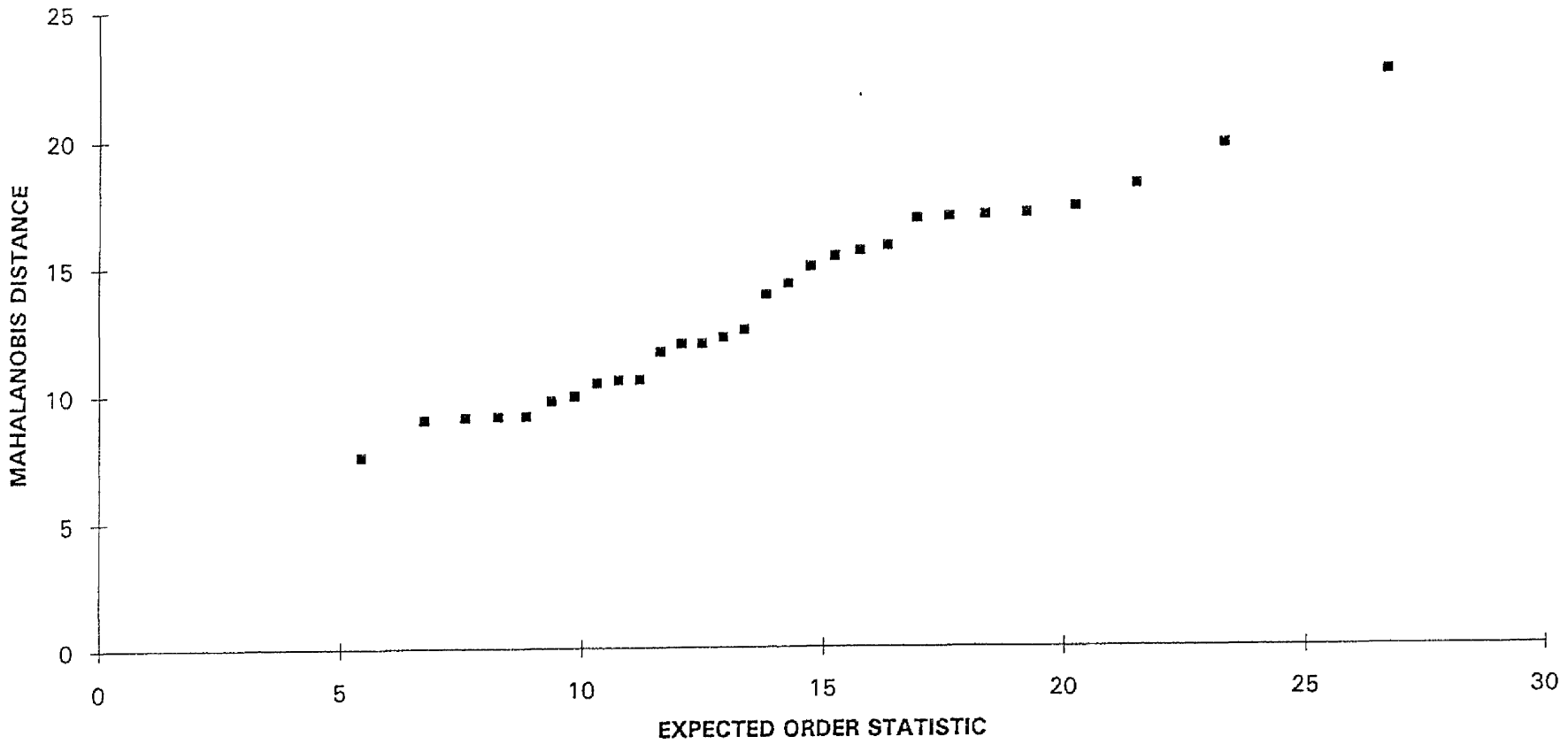
An important assumption behind each of these analyses is that the data is normally distributed. In the MANOVA analysis, it is assumed that $\underline{X}_{ij}$ - MVN(μ_j,Σ_j). In order to check this assumption probability-plots of mahalanobis distances (Gnanadesikan, 1977) were constructed for each group. These are shown in Figure 4.2a,b.

The data in both groups appear to satisfy the assumption of multivariate normality reasonably well.

4.5.2 Equality of Covariance Matrices

Equality of covariance matrices is another assumption underlying both these analyses. Box's M test (1949) is often used to test the hypothesis H: $\Sigma_1=\Sigma_2(=\Sigma)$. For the fourteen variable data, the test is very highly significant (M=196.1; P=0.001). Thus, there is evidence that covariance matrices differ. Likewise, the test for seven dimension matrices of hemispheric sums and differences are also highly significant (M=47.7; P=0.035) and (M=78.7; P=0.000).

FIGURE 4.2A - Probability-Plot For Normal Group



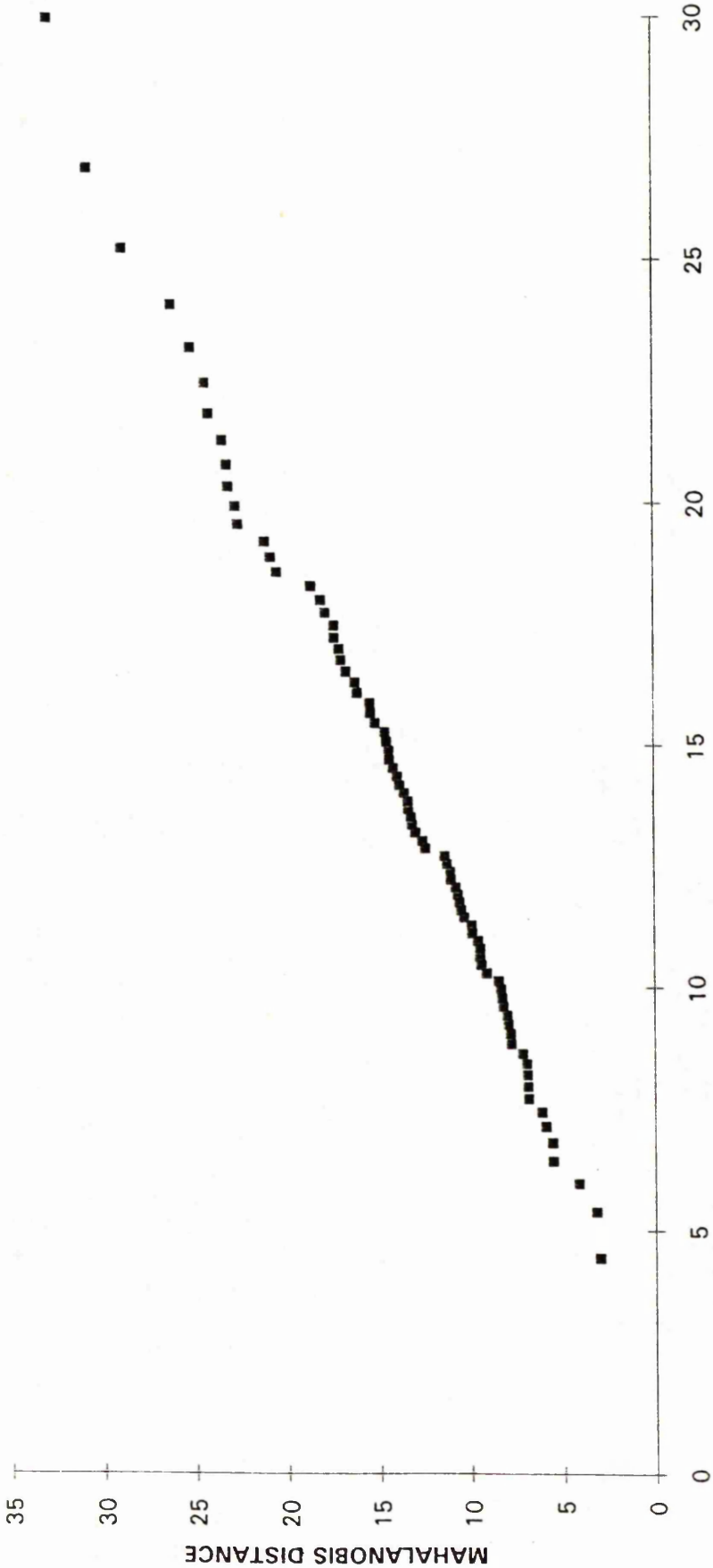


FIGURE 4.2b - Probability-Plot For DAT Group

The inequality of these matrices clearly complicates the results of all the preceding analysis. For one thing it goes some way to explain the non-significant result from the Hotelling's T^2 test applied to the 14 dimension mean vectors since it is known that the power of this test can be quite low when covariance matrices are not equal and group sample sizes differ (Davidson, 1972) - as is the case with our data.

4.5.3 Choice of Scale for Analysis

Thus far we have looked at square root transformed ROI data. The justification for this scale over others is now given. In the models presented thus far it was assumed that terms enter additively. This is based on the belief that a simple mathematical form can be used to describe the structure among regional means. In previous studies, the choice appears to be somewhat arbitrary, with some analysis based on raw data (Tyler et al, 1988) while others have used logged data (Worsley et al, 1991).

Suppose we reconsider the univariate model for the single group case. Thus, we assume a model of the form

$$X_{ij} = \mu_j + P_i + \varepsilon_{ij}$$

where μ_j is the mean response in the jth region, P_i is a random effect due to subject i and ε_{ij} is the residual error term. We further assume that P_i - $N(0,\sigma^2)$ and e_{ij} - $N(0,\tau^2)$. Now this is a basically very simple additive

model. Consequently, it may be seen as too simple to describe data on such a complex organ as the brain. More complex models of these datasets have been proposed including the scaled subprofile model of Moeller et al (1990).

For our datasets we studied the adequacy of this model by studying the effect of different transformations of the ROI data. In order to analytically identify the most suitable scale we used Tukey's 1 d.f. test for non-additivity (Snedecor & Cochran, 1980), regarding our data as a row (subjects) and column (regions) design. We considered transformations of the form $\underline{Y}=\underline{X}^{\lambda}$ for λ =0.0,0.25,0.5,0.75, 1.0. This includes $\log_e$ (λ=0.0), square root (λ=0.5) and raw untransformed mean counts data (λ=1.0). This was done for the normal and DAT groups separately. The results are summarised in table 4.5 in terms of the significance of the 1 d.f. F-test for non-additivity.

The best transformation is taken to be the case where the F-test is not significant. This is apparent from the p-values in tables 4.5(i), (ii). In both groups a range of values seemed reasonable, with the square root (λ=0.5) transformation the best of those considered. Both the raw untransformed and logged data indicated the presence of a significant non-additive component to error terms.

In order to assess whether this had a substantial effect on the region and region x group F-tests the RM ANOVA was carried out on log (λ=0.0),

λ	Non-Additive Error MS	Remaining Error MS	F-Ratio Non-Additivity	P-Value
0.0	0.032	0.0045	7.279	0.0073
0.25	0.005	0.0034	1.467	0.2266
0.5	0.0001	0.1742	0.0008	0.9771
0.75	5.881	5.0228	1.1709	0.2799
1.0	464.638	116.1315	4.0009	0.0462

(i) Normals

λ	Non-Additive Error MS	Remaining Error MS	F-Ratio Non-Additivity	P-Value
0.0	0.0724	0.0083	8.720	0.0032
0.25	0.0092	0.0062	1.489	0.2226
0.5	0.0700	0.3019	0.232	0.6304
0.75	38.4697	8.5110	4.519	0.0337
1.0	2699.384	193.8987	13.922	0.0002

(ii) DAT

Table 4.5 Tukey's 1 d.f. test for additivity. Look-up degrees of freedom in normal group F(1,363) and in DAT group F(1,1013).

square root (λ=0.5) and raw (λ=1.0) data. The results of these analyses are given in Table 4.6. It would appear, from the similarity of F-ratios in each row, that the effect of scale is only marginal. Thus, at least for these data, the scale is relatively unimportant to the outcome of the tests. Whether this would be the case in other datasets i.e with smaller sample sizes and/or higher dimensional data sets, is unknown.

		λ	
TEST	0.0	0.5	1.0
Region	37.7	39.1	37.9
	(0.0000)	(0.0000)	(0.0000)
Group			
x Region	3.22	3.44	3.40
	(0.0001)	(0.0000)	(0.0000)

Table 4.6 Univariate F-Ratios (unadjusted P-Values) for transformed datasets.

4.6 DISCUSSION

In this chapter we have looked at the use of some repeated measures ANOVA techniques for comparing regional mean profiles. As imaging datasets are generally

characterised by high dimensional/low sample sizes, it seems reasonable to assume that the univariate approach will be the most often used. Indeed, in some situations the univariate approach will be the only one possible because sample covariance matrices will be singular.

Even when there are more subjects than regions (as here) the MANOVA approach may lack the power of the univariate approach to detect differences. This is especially the case when covariance matrices are not equal and one group is larger or substantially larger than the other (Davidson, 1972). This possibly explains why the univariate ANOVA gives a significant difference between groups while the MANOVA does not. A further demonstration of the sensitivity of the univariate approach over the multivariate approach is obtained by carrying out the global tests for group x region interaction on hemispheric sums, this time excluding the occipital regions. In the univariate test the region x group test is still significant (F=3.50 ;UNAD-p=0.0040, GG-p=0.0160, HH-p=0.0146) while the MANOVA approach is not significantly different (T^2=7.674 ;p=0.3030). One other aspect of the univariate RM approach is that other explanatory factors e.g.sex or disease subgroups can be easily included into the analysis.

For follow-up analysis the variable selection approach based was useful although several runs may be necessary to fully understand differences between groups. Other methods here include the subset selection procedures described by Lehmacher (1990).

In the the choice of scale analysis we used Tukey's test which is not strictly designed to be applied to related columns (regions) data. A plus with this approach, if it is valid, is that it may be used when there are more regions than subjects. In our example the different transformations did not materially affect the results of the F-tests. This would not necessarily be the case in samll studies where the data contains outliers. Here appropriate transformations may lessen the impact of the outliers .

CHAPTER 5

INVESTIGATING THE INTERRELATIONSHIPS AMONG REGIONS

5.1 INTRODUCTION

The interest shown by brain researchers in imaging datasets is based on a biological model relating the status of the various cerebral processes to human behavioural and intellectual function. In these terms, it might seem reasonable to assume that the nature of this biological linkage would vary from the normal state to the diseased state; one manifestation in the diseased state being changes in behaviour or in the ability to perform simple intellectual tasks which could have been done before the onset of any disease. The effort of much of the research in this area is therefore an attempt to trace the clinical symptoms, characterising a disease, back to the functional source e.g. blood flow or glucose utilisation. It is no wonder then that investigators have used functional imaging datasets to study the functional association among brain regions in both normal and diseased states (Clark et al, 1984; Horwitz et al, 1984; Metter et al 1984a,b; Bartlett et al,1987; Worsley et al, 1991) and between different cerebral variables e.g. glucose and blood flow in the same brain regions (Kelly et al, 1982). The usual approach is to identify and scan a group of individuals with common clinical symptoms and then to study the correlation structure in their ROI data.

In this chapter we will look at the use of correlation analysis to investigate the nature of functional association among regions. A testing procedure, which can be used to identify differences in correlation structures among disease groups, is described. Multivariate techniques for exploring the underlying structure of imaging datasets including an approach to dimension reduction are investigated. In the last two sections, patterned covariance structures of relevance to the study of these datasets are reviewed, one of which is described in some detail.

5.2 INTERPRETING CORRELATIONS

An obvious measure of association between the levels of cerebral activity in two regions, and by implication, functional association between them, is the sample correlation coefficient. We will start this section by reconsidering the correlation matrix in Table 3.3. This gives the sample correlation coefficients between regional measurements on the raw data (albeit hemispheric sums). In other contexts, the occurrence of such large positive values might be taken to imply strong association between the underlying variables. However, in this context, this simple view cannot be taken so readily. Recall from section 3.2 that SPECT imaging datasets are dominated by large scalar differences resulting in substantial between subject variation.

5.2.1 Adjusted Data and Correlation Analysis

In common with the univariate analysis of means, the standard approach to studying possible functional associations is to adjust the data before calculating correlations. Tables 3.4(a)-(c) show the sample correlation coefficients following three such methods of adjustment. These involve normalisation to a single region 3.4(a), a subject mean 3.4(b) and through partial correlations 3.4(c) -adjusting for the subject mean. This last form of adjustment has been used most often (Horwitz et al, 1984, 1987; Soncrant et al, 1986). It is a more sophisticated attempt to study the relationships between regions by modelling the subject mean as a covariate.

Visual inspection of the correlation matrices in Table 3.4 suggest differences between the three approaches. It is noticeable, however, that the partial correlations are similar to those obtained by normalising to the subject mean.

5.2.2 Statistical Significance v's Functional Association

The statistical significance of these correlations might be formally tested using univariate methods as per Clark et al, (1984) Horwitz et al, (1984), Metter et al (1984a,b) and Bartlett et al, (1987). This would involve simultaneous testing of the hypothesis of zero correlation between all pairs of regions. This hypothesis

was tested for each coefficient in the sample correlation matrices given in table 3.4a and b. In order to test the null hypothesis we need to refer to the t-distribution on n-2=27 degrees of freedom. For two-sided tests with individual 5% significance levels, the null hypothesis should be rejected for an observed correlation coefficient greater than +.311 or less than -.311. With so many tests it is of course often advisable to control the overall type I error. Thus, a Bonferroni correction for each individual test was applied. For the correlations in table 3.4, this meant using individual _ significance levels of 5/21% and 5/15% for table 3.4b. Thus in 3.4a we reject the null hypothesis where the sample correlation coefficients are greater than +.542 or less than -.542 and in 3.4b where we have sample correlations greater than +.527 or less than -.527.

A method of displaying the results of these analysis is as a network of regions (with each region suitably identified) and with significant correlations identified by regions connected by lines (Bartlett et al, 1987; Horwitz et al, 1990). Thus, for these data the significant correlations (after Bonferroni adjustment) in each correlation matrix are illustrated in figure 5.1(a) & (b). Note, that the chosen arrangement of regions in these plots does not conform to any statistically derived configuration; it is merely a convenient graphical arrangement.

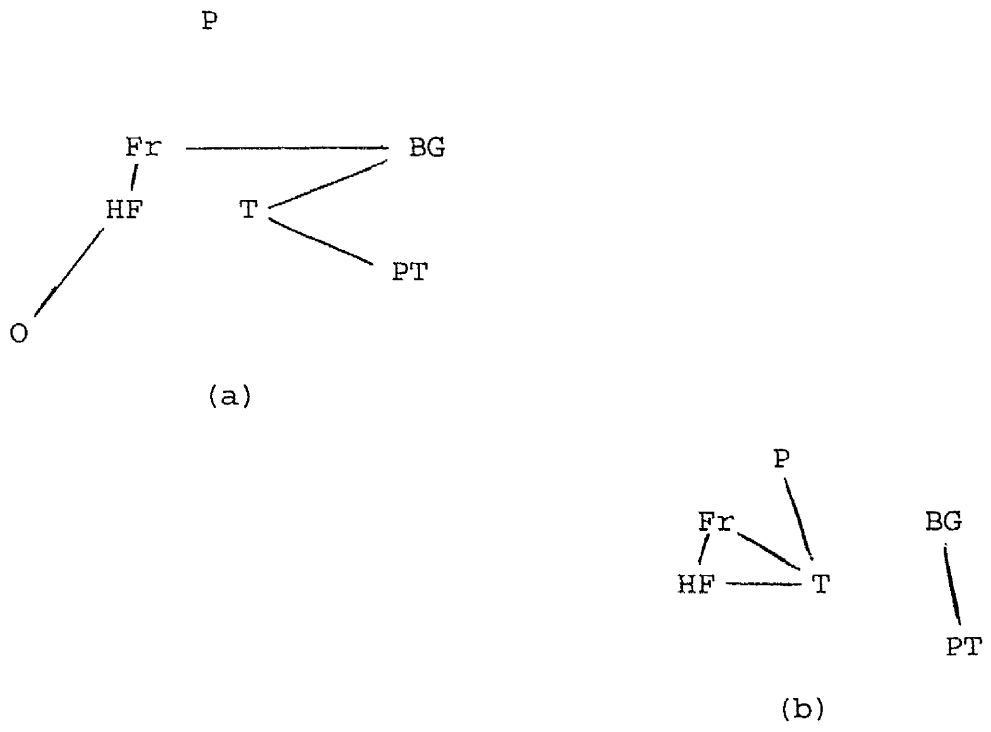


Fig 5.1 Representation of statistical significance of correlation coefficients between regions, based on (a) data normalised to subject mean and (b) data normalised to occipital activity.

The pattern of functional association among normalised regional variates seems are quite different in each of these figures. However, it should be remembered that the data are adjusted differently in each analysis. As mentioned previously, the interpretation must be made in terms of normalised variates and not in terms of the original regional variates. Such a mistake could easily

lead the unsuspecting analyst to two very different biological interpretations about such functional association. As in the between group analysis of means in section 3.4, it would seem therefore that simple conclusions cannot be extracted from the analysis of adjusted data. This is a major problem associated with data from cross-sectional studies, where we only have one vector per subject. Attempts at a formal study of functional association are confused by the need to adjust the data. The bottom line is therefore, that it is difficult to infer biological conclusions about functional association from these simple correlation analysis.

This is not to say that correlations themselves are of no value. Valuable insight into diseases may still be gained in between group analysis of correlation structures.

5.2.3 Problems of Interpretation

It can be demonstrated mathematically how correlation analysis, of adjusted datasets, can induce spurious large correlations among variables and hence compromise subsequent interpretation of functional association between regions. Consider the data normalised to occipital activity, as in Table 3.4(b). Pearson (1897), provides a formula for the correlation of two variables X_i and X_j expressed as ratios to a (common) third variable X_k, in the form

$$\text{cor}(X_i/X_k, X_j/X_k) = \frac{CV_k^2}{\text{sqrt}((CV_i^2+CV_k^2)(CV_j^2+CV_k^2))}$$

where CV_i is the coefficient of variation for the variable X_i.

This means that, in the case where, say, $CV_1=CV_2=CV_3(=CV)$ and X_i, X_j and X_k are independent, the correlation $Cor(X_i/X_k, X_j/X_k)=0.5$. In effect, we obtain a positively biased estimate of the correlation between variables i & j simply from involving a third variable in this way in the calculations. Note however that in the case of our data the X_i's are strongly correlated.

Similar difficulties arise when regions are normalised to a global subject mean. This is because normalisation introduces a sum constraint into each subject's data vector, of the form $\Sigma P_{j=1}\ Y_{ij} = p$. The normalised data is now of the compositional form (Aitchison, 1986). This introduces singularity into covariance and correlation matrices which as Aitchison pointed out leads to problems of interpretation. A further problem of interpretation, with this form of normalisation, could likewise arise if different subsets of regions were analysed. In the case of normalisation to a subject region vector mean, different inter-regional correlation matrices might result. This is similar to difficulties of interpretation found in partial subcomposition correlation analysis of compositional datasets (Aitchison, 1986).

One way around this would be to normalise the data to a global subject mean i.e. the slice average. However, the sum of normalised data may still be close to a constant in all subjects since the global slice mean will be strongly correlated with the subject vector mean. Near singularity may still be a problem with covariance matrices. After all, the global mean is effectively the average of all p-dimensional ROI data and an unobserved variable X_{p+1} - the complement of $\underline{X}$. And the problem of unclear interpretation would still exist.

Similar difficulties arising from the results of partial correlation analysis have been demonstrated. Ford (1986) has shown that, where the true correlation structure exhibits regional clustering, the observed partial correlations may be quite different from the expected correlations.

5.3 COMPARING CORRELATIONS BETWEEN GROUPS

Commonly, if we wish to compare the sample correlation matrices between different groups, we resort to the use of Box's M-test (1949). This is a likelihood ratio test (LRT) (Anderson, 1958) assuming the data follow multivariate normal distribution. In the two group case this provides a test of the hypothesis H: $\Sigma_1=\Sigma_2(=\Sigma)$ against the alternative of inequality.

For many statistical problems e.g. testing homogeneity of covariance assumptions in MANOVA and RM ANOVA, this is a useful test. However, the problem with

the test per se, is that rejection of the null hypothesis does not automatically imply inequality in the corresponding correlation matrices. For example, in the situation where covariance matrices are proportional i.e. $\Sigma_2 = c\Sigma_1$, the equality of correlations still holds even though overall equality of covariances does not. Other situations can be imagined where correlations are similar even when differences in variances and covariances exist between groups.

If the LRT is applicable (n>p) and the null hypothesis of equality of covariance matrices is rejected, a useful follow-up analysis is given by Manly and Rayner (1987). This involves a hierarchical testing procedure which partitions the overall LRT test statistic into three chi-square nested sub-tests. Each of the null and alternative hypothesis are composite assuming a different form for the difference in covariance structure between groups. As with the overall LRT, each of the subtests are based on the assumption of multivariate normality. The null and alternative hypotheses in this testing scheme are given below. The hierarchical test procedure is as follows. Test 1 is carried out first. If the null hypothesis is rejected then we conclude that the covariance and correlation matrices are different. If we cannot reject the null hypothesis then there is insufficient evidence for differences in correlations between groups and proceed to Test 2. If the null hypothesis is rejected we conclude that correlations are not significantly different but some or all of the

regional variances are different between groups. If we cannot reject this hypothesis then we perform Test 3. This provides a test of the null hypothesis H_0: $\Sigma_2 = \Sigma_1$ against the alternative H_1: $\Sigma_2 = c\Sigma_1$. Note, that the comparison of H_0 in test 3 with H_A in test 1 is just the overall likelihood ratio.

As we saw in section 4.4.2, the LRT strongly rejected the hypothesis of equality of covariance matrices both for hemispheric sums and differences. We will therefore use the approach described above to investigate these differences further. The maximised loglikelihoods (less the common constant term), test statistics and p-values for each sub-test on the sums and differences separately are given in Table 5.1(a), (b). Since the overall LRT was assessed at the 5% level, the significance of each subtest is made at the 5/3=1.7% level - as recommended by Rayner et al (1990).

Unfortunately, not all log likelihoods were found for the hemispheric sums (Table 5.1(a). The log likelihoods l_1 and l_2 are obtained by iteration using fixed point or one point approaches (Atkinson, 1978). Unfortunately, the method described by Manly and Rayner in their paper, doesn't always converge in the Fortran program I have written to calculate the log likelihoods. Manly and Rayner suggest that in their experience l_1 and l_2 always converge although they do not give any proof, or guidelines as to any special steps in their computer program.

Test	Null Hypothesis	Alternative Hypotheses
1	Equal Correlations $\Sigma_2 = C\Sigma_1 C$, where $C=diag(c_{11},\ldots c_{1p})$	Unequal Covariance Matrices $\Sigma_2 \not= \Sigma_1$
2	Proportional Covariance Matrices $\Sigma_2 = c\Sigma_1$	Proportional Variances $\Sigma_2 = C\Sigma_1 C$, where $C=diag(c_{11},\ldots c_{1p})$
3	Equal Matrices $\Sigma_2 = \Sigma_1$	Proportional Matrices $\Sigma_2 = c\Sigma_1$

Consequently, for these data, not all test statistics are given. See Table 5.1(a). This is unfortunate since T_3 is an important test for us. The comparison of l_3 with l_1 gives a likelihood ratio statistic $2(l_3-l_1)=37.82$, which when compared to a chi-square distribution on 27 degrees of freedom gives a p-value of 0.0807. This is not significant at the 5 or 1.7% levels. Test 1 is highly significant. Hence, there is evidence that the elements in the group covariance matrices differ by a proportionality constant.

Happily, all log likelihoods were calculated for hemispheric differences. Hence test statistics were

(a) Hemispheric Sums

Model Cov. Form	Log Likelihoods	Test	Test Statistics	D.F.	P-Value
Σ_2, Σ_1	l_3 = -50.62				
		1	T_3 = ?	21	?
$\Sigma_2 = C\Sigma_1 C$	l_2 = ?				
		2	T_2 = ?	7	?
$\Sigma_2 = c\Sigma_1$	l_1 = -69.53				
		3	T_1 = 9.93	1	0.0016
$\Sigma_2=\Sigma_1 (=\Sigma_0)$	l_0 = -74.50				

(b) Hemispheric differences

Model Cov. Form	Log Likelihoods	Test	Test Statistics	D.F.	P-Value
Σ_2, Σ_1	l_3 = 795.86				
		1	T_3 = 35.8	21	0.0232
$\Sigma_2 = C\Sigma_1 C$	l_2 = 777.99				
		2	T_2 = 15.0	7	0.0357
$\Sigma_2 = c\Sigma_1$	l_1 = 770.48				
		3	T_1 = 14.1	1	0.0002
$\Sigma_2=\Sigma_1 (=\Sigma_0)$	l_0 = 763.42				

Table 5.1 Hierarchical Tests on Equality of Covariance Matrices for (a) Hemispheric Sums and (b) Hemispheric Differences.

calculated and given in Table 5.1 (b). At the 1.7% level, we cannot reject the hypothesis of equal correlations although it is very close. Nor can we reject the hypothesis of proportional variances again very close. The only significant test is the one for proportionality of covariance matrices. Hence, the most plausible conclusion from comparison of covariance matrices is that they are proportional. In this case, the sample covariance for the DAT group is proportionately greater than for the normal group.

It should be noted that the tests described here can suffer from the same problems found in the overall LRT. Namely, the sensitivity of the tests to non-normality and the lack of power (Davidson, 1972) resulting from big differences in group sample sizes.

5.4 EXPLORATORY DATA ANALYSIS AND DIMENSION REDUCTION

5.4.1 Principal Components Analysis

Many investigators have studied inter-regional correlation structures using principal components type analysis (PCA) including the factor analytic approaches of Volkow et al (1986), Clark et al (1985a) and Moeller et al (1987).

Principal components analysis (PCA) is a technique for identifying the patterns in a multivariate dataset describing common sources of variation in a group of subject vectors. This involves taking an orthogonal

linear transformation of the data $Y=Q^TX$ such that $Cov(Y)=Q\Sigma Q^T=V$, where $Cov(X)=\Sigma$, V is the $diag(v_1,v_2,\ldots,v_p)$ of eigenvalues and Q has rows corresponding to eigenvectors $\underline{q}^T_j$, $j=1,\ldots,p$. Each transformation $y_j=\underline{q}_j^T\underline{x}$ is termed a principal component (PC) and involves a linear combinations of the original data such that the variance $Var(y_j)=v_j$. The larger v_j is, the larger is the variation in y_j among subjects. Thus, the y_j with the largest variance describes the single most common source of variation among subjects and is described by the eigenvector $\underline{q}_j$.

As an illustration of the use of this technique, a PCA was carried out on our normal group based on the sample covariance matrix of (unnormalised) square root data. All fourteen regions are used. Table 5.2a shows the first four eigenvalues and eigenvectors corresponding to the first four PCs. Not surprisingly, the coefficients of the first PC are all of the same sign and approximately the same magnitude. This amounts therefore to a subject average, of all regions, which accounts for 94.7% of the total variation. This reiterates the extent to which the large between subject scalar effect dominates these datasets. From the sign and magnitude of the coefficients of the second PC we see that this describes a contrast between a grouping of right and left occipital & parietal regions with right and left basal-ganglia. Although this accounts for only 2.1% of the overall variation, it is about 2.1/5.3 x 100=39.6% of variation in patterns among subjects. Interpreting the other PC's in a similar

fashion we see that PC_3 accounts for a further 1.3/5.3 x 100=24.5%.

For comparison, PCA was also carried out on the DAT group. Table 5.2b gives the first four eigenvalues and eigenvectors from PCA of the covariance matrix. Not surprisingly, PC_1 again describes a subject average. PC_2 accounts for 2.8/8.6 x 100=32.5% and describes a contrast between right and left parietal, posterior-temporal and occipital regions with right and left frontal, higher-frontal and basal-ganglia regions. Interestingly, the two groupings of regions, described by this contrast, correspond roughly to a comparison between levels of regional activity in the posterior area of the brain with the anterior area of the brain.

Since, in each of these analyses, PC_2 appears to involve a different contrast among variables, we might tentatively conclude that the largest important source (described by each PC2) of variation, after PC_1, among subjects is different in each of the groups. However, this is not entirely conclusive or based on exactly the same number of variables. Nor has it been subjected to any form of statistical test; rather, it is a judgemental interpretation.

One common feature which emerges in each of these PCA analyses has been the close link between bilateral regions (right and left). With few exceptions, the coefficients for right and left regional variables have been similar in sign and magnitude. Thus, the suggestion of working with hemispheric sums and differences, may not

REGION	PC_1	PC_2	PC_3	PC_4
Right Frontal	-0.27	-0.05	0.18	-0.34
Left Frontal	-0.28	0.01	0.16	-0.21
Right Hi-Frontal	-0.26	0.03	0.28	-0.15
Left Hi-Frontal	-0.27	0.02	0.15	-0.19
Right Temporal	-0.27	-0.05	0.40	0.36
Left Temporal	-0.29	0.06	0.23	0.29
Right Parietal	-0.25	0.31	0.22	0.03
Left Parietal	-0.26	0.28	-0.05	-0.30
Right Post-Temporal	-0.26	0.09	-0.15	0.16
Left Post-Temporal	-0.26	0.15	-0.44	-0.44
Right Basal-Ganglia	-0.28	-0.63	0.02	0.06
Left Basal-Ganglia	-0.29	-0.52	-0.38	0.02
Right Occipital	-0.24	0.22	-0.32	0.40
Left Occipital	-0.25	0.25	-0.34	0.31
% Variance	94.7	2.1	1.3	0.6
Cumulative	94.7	96.8	98.1	98.7

Table 5.2a PCA of covariance matrix for Normal Group based on unnormalised data. Components roughly describe contrasts:

PC1 - Weighted average of all regions.

PC2 - (RP, LP, RO, LO) v's (RBG, LBG)

PC3 - (RF,LF,RHF,LHF,RT,LT,RP) v's (LPT,LBG,RO,LO)

PC4 - (RT, LT, RO, LO) v's (RF, LF, RHF, LHF, LP, LPT)

REGION	PC_1	PC_2	PC_3	PC_4
Right Frontal	-0.26	-0.15	-0.23	-0.23
Left Frontal	-0.27	-0.18	-0.26	-0.30
Right Hi-Frontal	-0.27	-0.20	-0.29	-0.05
Left Hi-Frontal	-0.27	-0.22	-0.31	-0.11
Right Temporal	-0.28	-0.04	-0.18	0.10
Left Temporal	-0.27	-0.06	-0.18	0.26
Right Parietal	-0.27	0.21	-0.10	0.10
Left Parietal	-0.26	0.19	-0.00	0.06
Right Post-Temporal	-0.26	0.38	-0.03	0.40
Left Post-Temporal	-0.25	0.30	0.07	0.36
Right Basal-Ganglia	-0.29	-0.32	0.49	0.11
Left Basal-Ganglia	-0.30	-0.42	0.52	0.06
Right Occipital	-0.25	0.34	0.24	-0.35
Left Occipital	-0.22	0.38	0.21	-0.56
% Variance	91.4	2.8	2.2	0.8
Cumulative	91.4	94.3	96.5	97.3

Table 5.2b PCA on sample covariance matrix for DAT Group based on unnormalised data. Components roughly describe contrasts:

PC1 - Weighted average of all regions.

PC2 - (RP, LP, RPT, LPT, RO, LO) v's (RF, LF, RHF, LHF, RBG, LBG)

PC3 - (RBG, LBG, RO, LO) v's (RF, LF, RHF, LHF, RT, LT)

PC4 - (LT, RPT, LPT) v's (RF, LF, RO, LO)

be unreasonable. This is investigated in a subsequent section.

5.4.2 Graphical Representations of Variation Among Subjects

As part of the exploration of any multivariate dataset it is common to seek useful graphical displays to highlight the outstanding features, for example extreme outlying cases, clustering of cases or variables. The profile plots at the beginning of chapters 3 and 4 were very useful in showing the hierarchical pattern in the levels of activity. A useful display of the variation among individuals, in terms of common pattern of activity, may be provided by PC analysis. Since the PCs describe linear combinations of the original variables describing the greatest sources of variation, it follows then that a useful summary plot of the variation among individuals is in terms of the scores on the first couple of PCs. This approach is used later to select individuals in an analysis of the correlation of the imaging data with other imaging datasets.

5.4.3 Graphical Representations of Association

To study the association between variables, a low dimensional summary display is useful. We have already seen, in figures 5.1a,b, an attempt to give graphical expression to the association between regions found from significance testing. The problem, of course, was that

the organisation of regions was totally arbitrary and gave no indication of the magnitude of correlations. In this section we will describe a statistical approach to the problem of representing association between regions in a non-arbitrary way.

First we will start by identifying a measure of similarity of some relevance in previous analyses. In section 4.4.1 we saw that the sphericity assumption was violated for the pooled sample covariance matrix, meaning that the variance of pairwise differences of variables i.e. $Var(X_i-X_j)$, was not constant for all pairs (i,j). In both groups, the sphericity assumption was very strongly rejected. We will now study the pattern in these variances further. The matrices of sample variances of all pairwise differences, in each group, are given in table 5.3. Since the variances are symmetric i.e. $Var(X_i-X_j)=Var(X_j-X_i)$ only one half of the matrix is given for each group. The Normal data are shown in the lower triangle and the DAT group in the upper triangle. To investigate the structure in these variances, multidimensional scaling techniques were used. $Var(X_i-X_j)$ is regarded as a dissimilarity measure between variables i & j and we attempt to fit a low dimensional model to describe the configuration of these dissimilarity measures. Nonmetric multidimensional scaling (MDS) (Kruskal, 1964), as provided in the SPSSX ALSCAL program (SPSS inc), was used to provide 2-dimensional displays of the configuration of variances in each group. A Fortran program was used to construct the pairwise difference

	RF	LF	RHF	LHF	RT	LT	RP	LP	RPT	LPT	RBG	LBG	RO	LO
RF		8.9	13.4	18.2	26.7	34.6	36.9	43.0	59.5	61.4	76.3	89.1	66.7	72.4
LF	4.8		20.1	13.1	30.5	33.3	53.5	43.3	74.3	62.9	86.1	89.6	76.0	79.3
RHF	13.0	11.9		11.0	21.4	36.4	34.4	47.2	64.9	71.6	79.7	90.2	77.2	95.1
LHF	11.9	7.4	3.6		25.4	34.6	50.5	47.2	73.6	68.8	84.5	91.9	82.9	97.2
RT	21.2	18.0	16.3	18.7		16.2	29.2	37.0	46.7	54.4	75.6	86.6	55.2	73.4
LT	24.4	15.1	17.0	13.4	10.3		47.6	41.9	54.6	54.5	87.2	89.4	70.3	82.4
RP	27.5	26.0	14.9	19.8	28.2	28.6		26.9	24.6	48.2	89.9	113.1	40.5	58.4
LP	25.6	22.2	17.2	13.8	37.6	25.2	14.6		38.7	26.0	88.0	95.4	41.0	46.6
RPT	21.4	16.8	23.8	18.4	24.7	20.9	23.9	20.0		27.8	109.8	133.6	39.9	55.4
LPT	32.2	30.9	38.9	30.3	58.1	44.7	42.9	18.7	18.4		100.3	115.0	44.7	51.1
RBG	43.9	48.1	51.8	48.9	48.8	56.9	91.1	85.1	58.8	77.8		39.1	93.5	119.7
LBG	55.2	52.6	58.7	47.1	63.5	58.9	92.8	71.5	51.5	60.6	27.9		111.8	137.6
RO	37.5	34.4	33.4	29.4	42.9	33.8	26.6	22.1	12.9	24.8	81.3	68.0		18.5
LO	37.4	31.8	34.5	29.7	47.6	36.7	27.3	22.9	17.1	24.6	86.6	70.7	3.2	

TABLE 5.3 Sample variance of pairwise differences $Var(X_i - X_j)$ x 100. Normal (lower triangle) and DAT (upper triangle)

FIGURE 5.2a 2-Dimension Solution For Normal Group

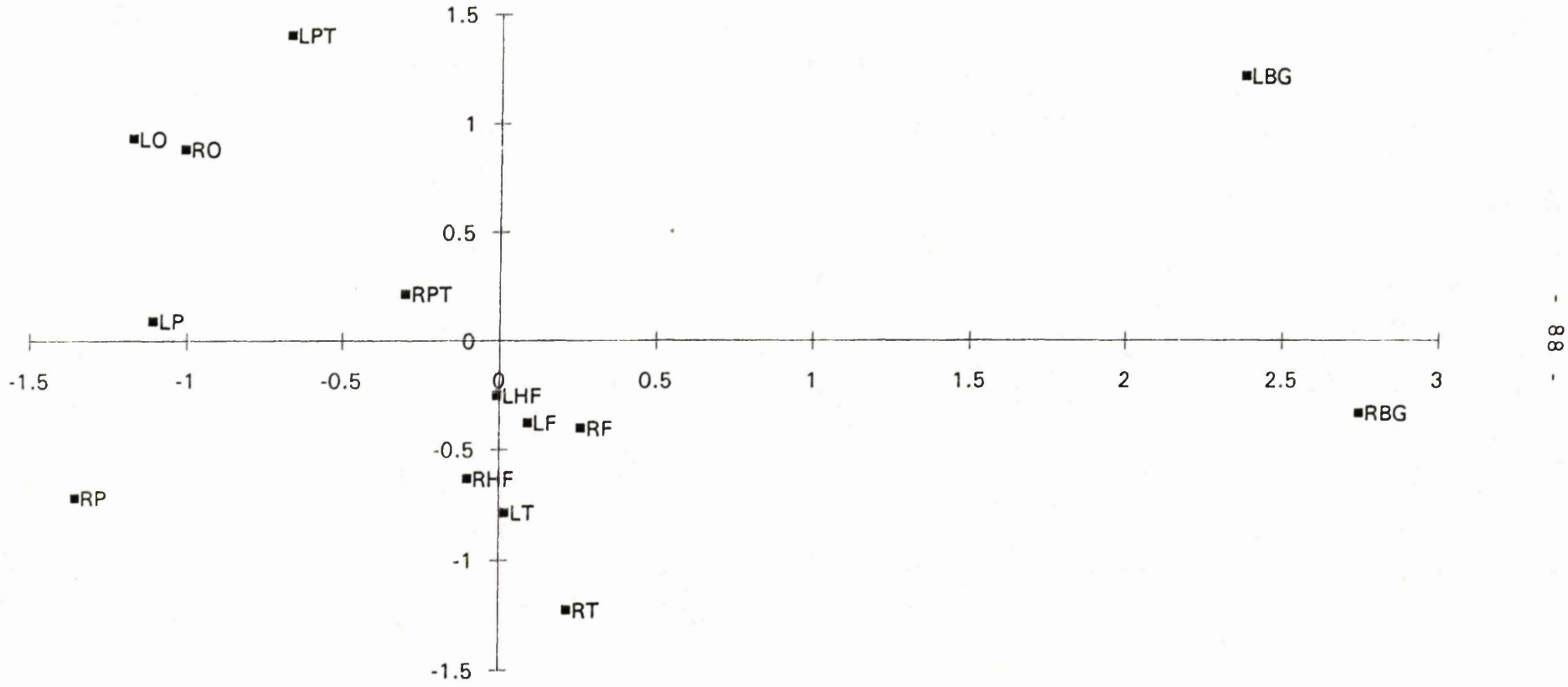
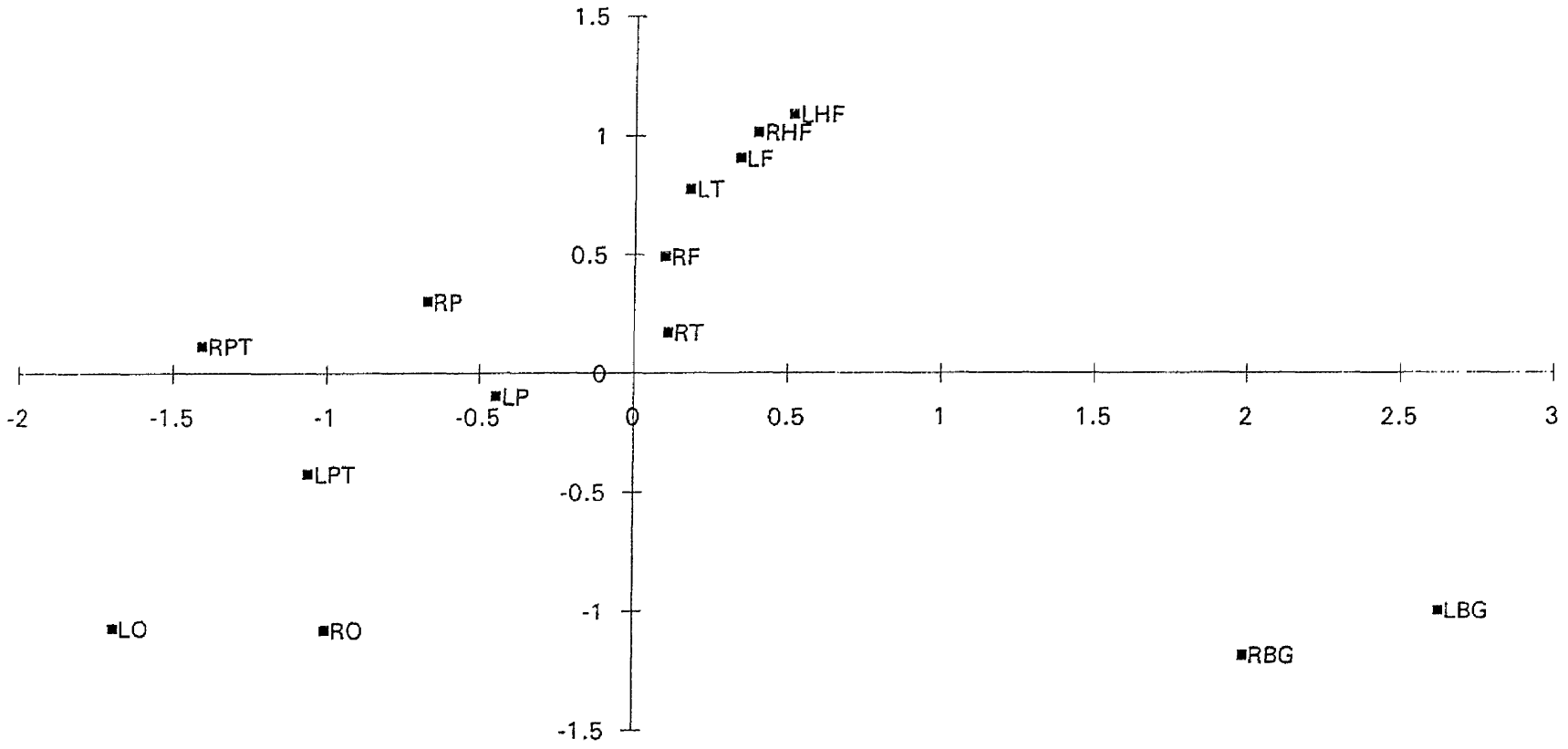


FIGURE 5.2b 2-Dimension Solution For DAT Group



variance matrix which was edited into an SPSSX command file. Two-dimensional solutions were fitted to each group and displayed in figures 5.2a,b. The most striking feature in both plots is the separation of the basal-ganglia region (right and left) from all other regions. This feature underlines an important fundamental difficulty in attempting to image some regions of the brain. The essential difficulty lies in the fact that the basal-ganglia is a smallish region - smaller anyway than the 13mm thickness of the slice (see section 2.5). Consequently this region can be difficult to identify reliably and hence associated data will be variable. In the figures all other variables appear roughly in alignment well separated from these regions.

Because of the sensitivity of MDS to individually large distances, regions close to each other will be less accurately fitted and hence some features of the observed configuration will be less reliably represented. In this example, there is little doubt that the basal-ganglia dominates figures 5.2a,b. In view of this the analysis was redone, excluding the basal-ganglia variable. The test of sphericity is still highly significant in each group even after excluding the basal-ganglia. The new MDS solutions are displayed in figure 5.3a,b. In both plots, most bilateral pairs are close indicating relatively small variances and probably larger correlations between regions. The exceptions are perhaps right and left parietal and posterior-temporal regions. In addition, each bilateral pair of regions are generally closer to

FIGURE 5.3a 2-Dimension Solution For Normal Group (Excl. BG)

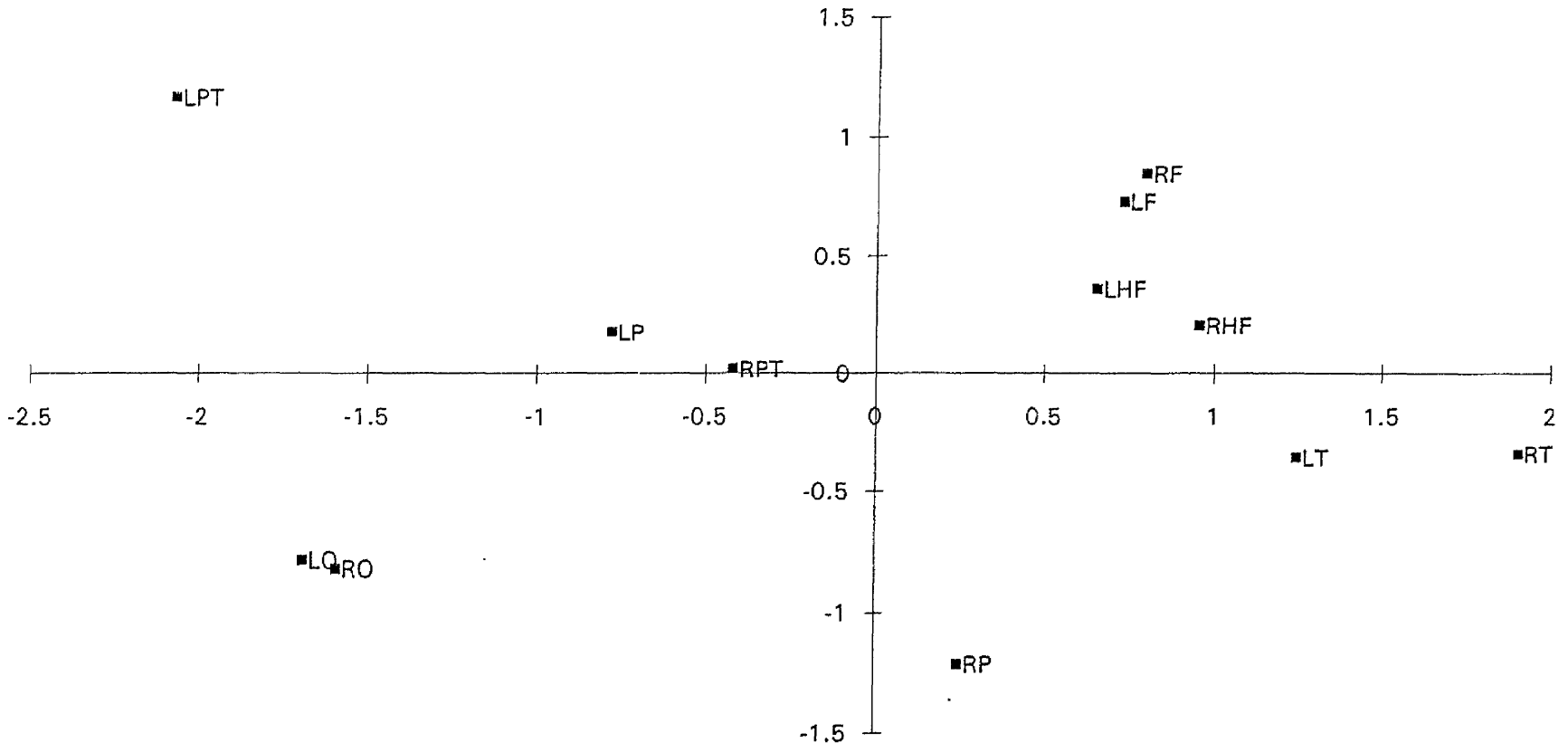
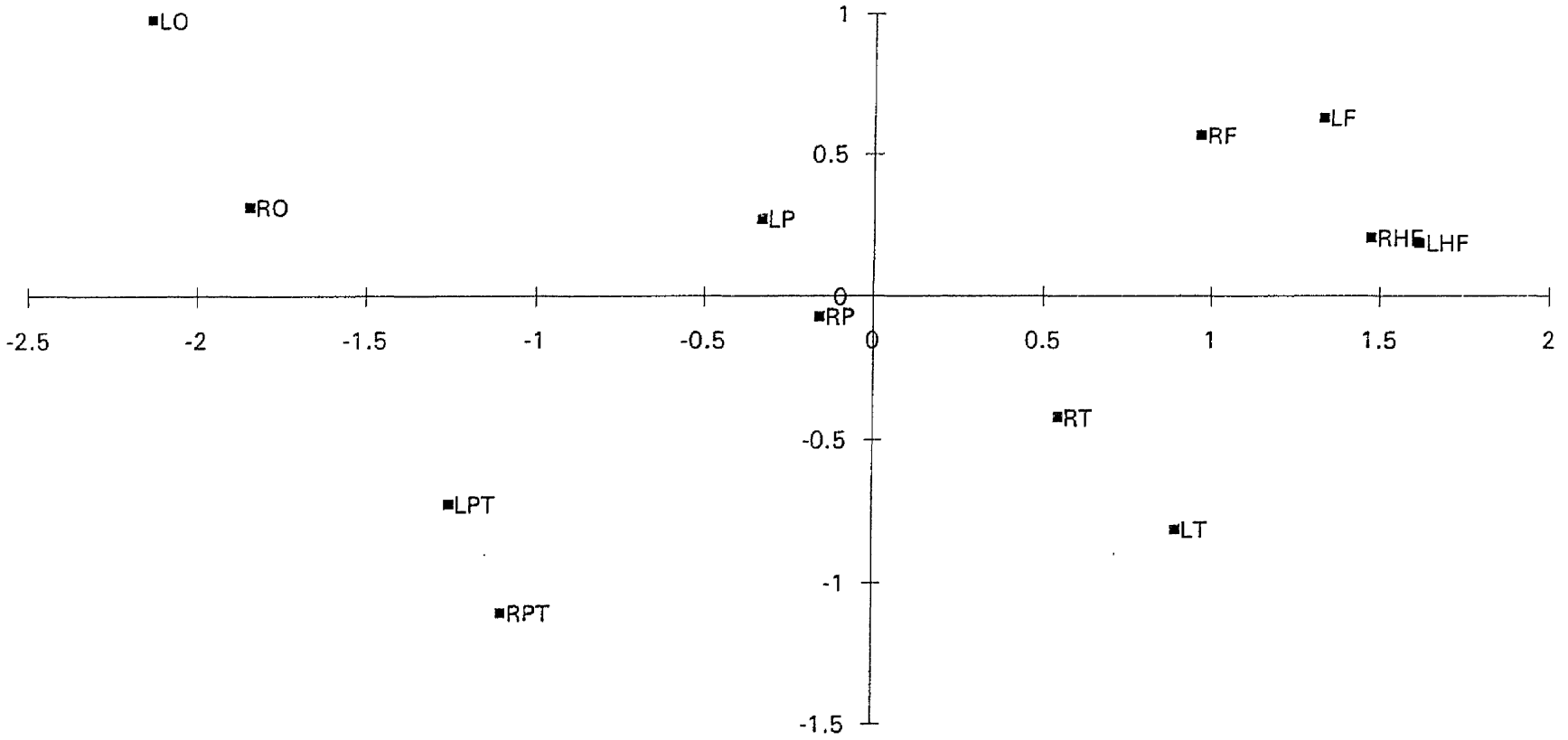


FIGURE 5.3b 2-Dimension Solution For DAT Group (Excl. BG)



each other than to other regions; with the exception of right and left parietal and posterior regions once again.

In the variance matrix (Table 5.3) it is clear that the larger pairwise variances are found in the DAT group. For example, the largest pairwise sample variance in the DAT group (after the basal-ganglia) is 0.972, between LO and LHF. This contrasts with the comparable sample variance in the normal group of 0.297. Since the the solutions are scaled differently it is not possible to compare the relative distances between any two regions between plots - only relative configurations.

It is interesting to note that, in each plot, the arrangement of regions along the x-axis correspond roughly with the positioning of regions along the anterior to posterior axis i.e. the long axis, of the brain. This tends to suggest that the variance of pairwise regional differences is related to the distance between regions.

Similar observations have been made by Worsley et al (1991) and Horwitz (1990). We will use this information as a basis for further study in a later section when an attempt is made to model the covariance structure.

5.3.4 Hemispheric Sums and Differences

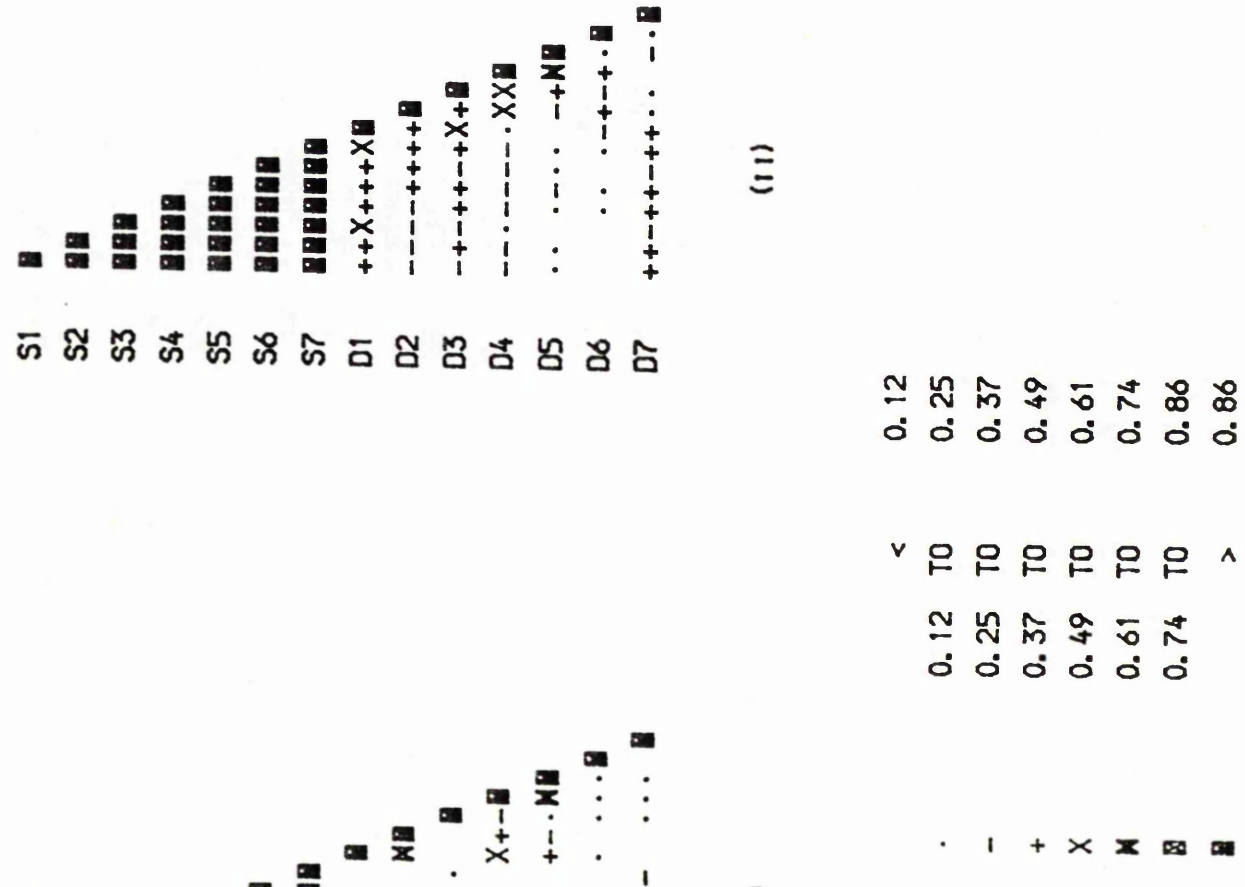
As mentioned in section 1.3, the ROI approach performs a dual role. On the one hand it is a means of summarising the pixel image in terms of comparable anatomical structures between different subjects. It also

helps reduce the dimensionality of the image dataset. Even so, we saw in table 1.1 that experimental costs often, as not, result in quite a large number of regions being measured in relatively few subjects. That is, the number of variables (p) can be close to, if not greater than, the number of subjects (n). Consequently, further dimension reduction may be required.

Worsley (1991) has recommended a form of dimension reduction which uses the fact that, often as not, we obtain measurements from the same region in both hemispheres e.g. right frontal and left frontal. He proposes that we should always transform this bilateral ROI data to hemispheric sums (Right+Left) and differences (Right-Left) and analyse these separately. This of course presupposes that sums and differences are independent of each other. For each group we calculated the bilateral sums and differences and calculated the correlations between them. These are presented in shade matrix form in Figure 5.4(i) & (ii). As expected for sums, the correlations are large. For the differences the correlations are less strong. For the sums and differences, the pattern and magnitude of the correlations seems different in each group. In the DAT group, only the occipital and temporal regions register any substantial correlation. In contrast, the normal group exhibit numerous correlations between sums and differences.

In order to assess the significance of these correlations, one possibility would be to carry out

Figure 5.4 Shade matrices for correlations between hemispheric sums and differences in (i) the DAT group and (ii) the normal group.



multiple tests of the hypothesis of zero correlation for each of the 49 pairings of sums and differences. Instead, an overall test of the independence hypothesis is obtained using the technique of canonical correlation analysis. Here we are looking for the linear combination of differences and of sums exhibiting the largest correlation. For the correlations described for the normal group in figure 5.4(i), the first canonical correlation is not significant at the 5% level ($p=0.2279$). Compare this with the equivalent test applied to the data in the DAT group. In this group the first canonical correlation is significant ($p=0.0143$). These results are perhaps somewhat surprising, considering the pattern of intensity illustrated in each of the shade matrices. The fact that in the DAT group, the test was significant is probably due in part to the large number of subjects involved. In saying that, we can see from figure 5.4(i), that the correlation is weak. Thus, little information will probably be lost by looking at sums and differences separately.

An interesting aside in this analysis is the effect of using different scales for the analysis. Here we have used square root data. If we carry out this analysis on the raw untransformed count data, almost identical results are obtained. For the normal group $p=0.2600$ and for the DAT group $p=0.0143$. However, if we use the log transformed data, the first canonical correlation is significant in the normal group at the 10% level ($p=0.0948$) and there are two significant canonical

correlations in the DAT group ($p=0.0014$, $p=0.0205$). The log scale suggests more association between sums and differences than the other two scales. Which result is correct is far from clear. However, the results earlier in section 4.5.3 would tend to suggest that square root data was the more appropriate scale for analysis of these data. Thus we might tentatively conclude that the log scale provides slightly less reliable conclusions about the association between sums and differences, than does the square root.

Another point worth noting about this analysis is due to the following result attributable to Pitman (1939) and Morgan (1939). The test of association between sums and differences between two variables is equivalent to a test of the equality of variances of the two correlated variables. We can see from the intensity of shading in figure 5.4(i) that only the occipital region has a sizeable correlation between sums and differences. By contrast figure 5.4(ii) shows a number of moderately large correlations between sums and differences - only basal ganglia being unshaded.

5.4 LINEAR SPATIAL CORRELATION MODEL

5.4.1 A Model Based on Distances Between Regions

As noted in section 5.3, there is some evidence to suggest that the variation in regional levels of activity is related to the distances between regions. Similar

observations on the correlation structure were noted by Worsley et al (1991) which prompted them to propose a model for the covariance structure of the form

$$\Sigma = a_1 J + a_2 D + a_3 I$$

where J is the unit matrix, $D = \{d_{ij}\}$ is a matrix of distances between all regions and I is the identity matrix. D is assumed known. Hence, the parameters to be estimated are a_1, a_2 & a_3, where a_1 corresponds to the between-subject variation, a_2 is the distance effect and a_3 the variation within regions. The number of unknowns in this model is three as compared to 1/2 x p x (p-1) in a full unpatterned covariance structure (see section 5.5.1).

The elements Σ_{ij}, of the covariance matrix are defined as

$$\Sigma_{ij} = \begin{cases} a_1 + a_2 d_{ij}, & \text{for } i \neq j \\ a_1 + a_3, & \text{for } i = j. \end{cases}$$

Covariances are expressed as a linear combination of the subject effect and an effect due to the distance between ROIs, while the variances involve a between subject and within region effect.

5.4.2 A Distance Matrix for SPECT

It would be interesting to investigate the spatial pattern in our SPECT data, especially as we have noted already, in section 5.3.3, a hint of a spatial effect.

In the paper by Worsley, the distance matrix D was found by matching the regions used in the study for data extraction with the distances (cms) in a standardised 3-dimensional atlas of the brain. Unfortunately, we do not have corresponding distances for the regions used in our SPECT experiments. We can however, construct a crude alternative using a nearest neighbour set of distances. The following scheme has been used. Adjacent regions are defined by touching boundaries as illustrated in figures 2.4a,b.

1. Different regions adjacent to each other and in the same hemisphere d=1.
2. Different regions adjacent to each other and in opposite hemispheres d=1.5.
3. Same region in different hemispheres d=0.5.
4. Regions with one region between them in same hemisphere d=2.
5. Likewise for two regions between them in same hemisphere d=3.
6. 4. And 5. above but in different hemispheres d=2.5 & 3.5.

The shortest notional distance between regions (using

this scheme) is used in this case.

5.4.3 Parameter Estimation

Worsley used maximum likelihood methods (Mardia & Marshall, 1984) to fit this model to data from a group of normal subjects, assuming regional variances (a_3) constant for all regions. Unfortunately, we do not have access to suitable programs for fitting this model to our datasets. However, note that the variance of the difference between regional variables can be written

$$^1/_2 \times V(X_i - X_j) = a_3 + a_2 d_{ij}.$$

This is a simple linear regression equation involving two of the parameters of interest to us - excluding a_1. A plot of the distances (given in the previous section) against sample variances of differences between variables (see table 5.3) in the normal group is given in Figure 5.5. There is visual evidence of association between these two measures and hence that the model may be quite reasonable.

Estimates of a_3 and a_2 can be obtained using ordinary least squares. However, because of the correlation between the $V(X_i - X_j)$'s, we cannot use simple least square standard errors. In order to obtain more approapriate estimates of standard errors I used a bootstrap approach (Efron, 1986). Random sampling was used to obtain bootstrap datasets X^* from our n x p

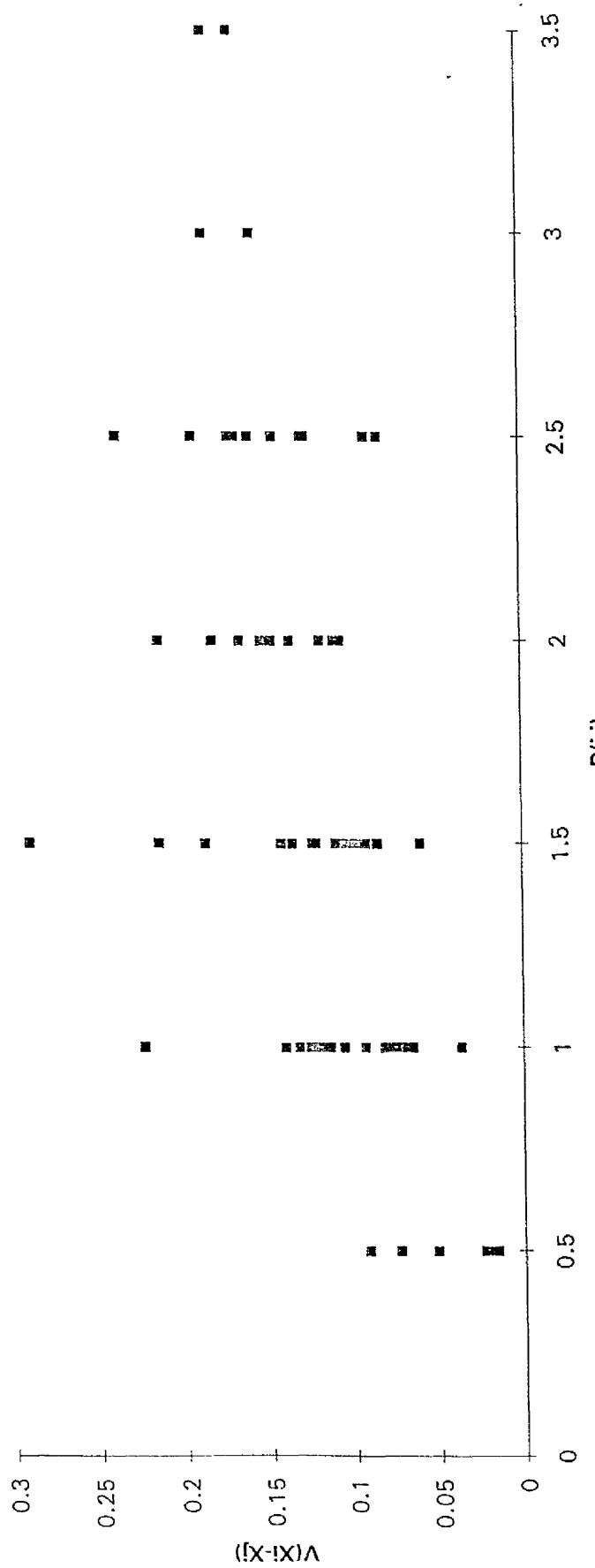


FIGURE 5.5 Plot of variance estimates of pairwise regional differences against notional distances between regions in SPECT

dataset X. For each X^*, pairwise sample variances of differences between variables $V(X_i-X_j)$ were calculated. These were then regressed on the distance matrix $D = \{d_{ij}\}$ described in section 5.4.2 to provide new estimates of a_2 and a_3, denoted by a_2^* and a_3^*. Two hundred such datasets were generated and from these bootstrap estimates were obtained. Standard errors, $SE(a_2)$ and $SE(a_3)$, were calculated from the sample variances among the bootstrapped a^*_2s and a^*_3s. These estimated parameters, for the normal and DAT groups separately, are given in Table 5.4.

		Normal Group			DAT Group		
Parameter	j	a_j ($\times 10^{-2}$)	$SE(a_j)$ ($\times 10^{-2}$)	T	a_j ($\times 10^{-4}$)	$SE(a_j)$ ($\times 10^{-4}$)	T
Distance	2	4.152	1.404	2.96	11.156	1.797	6.21
Region Variance	3	5.103	1.836	2.78	5.381	1.877	2.87

Table 5.4 Bootstrap estimates in normal and DAT groups.

5.4.4 Hypothesis Testing

The statistical significance of each parameter in the model can be assessed using the approximate t-

statistic $t_j=a_j/SE(a_j)$, for $j=2,3$. This can be compared to a value of 2 for deciding on the significance of the parameter. The observed t-statistics are given in Table 5.4. In both groups, both parameters are large enough to be considered significant. Thus there is evidence in support of a significant spatial effect.

As Worsley provided the data used in his research we may assess the reliability of the bootstrap approach used here. Bootstrap estimates of a_2 and a_3 and test statistics are given along with Worsley's actual estimates in Table 5.5. Although the t-statistics were slightly smaller with the bootstrap approach, I would say that they compare satisfactorily with the corresponding statistics obtained using maximum likelihood.

		Bootstrap Estimates			Worsley Estimates		Parameter
	j	a_j $(x10^{-2})$	$SE(a_j)$ $(x10^{-2})$	T	a_j $(x10^{-4})$	$SE(a_j)$ $(x10^{-4})$	T
Distance	2	.6367	.2235	2.84	.4325	.1304	3.32
Region Error	3	28.1	3.2	8.92	27.48	2.71	10.14

Table 5.5 Bootstrap estimates and actual parameter estimates using PET Data (Worsley, 1991).

5.5 COVARIANCE MATRICES RELEVANT TO IMAGING PROBLEMS

To follow on from the previous section we now review some models for covariance structures which have been considered in the literature, either to describe covariance patterns or to assess theoretically or by simulation the properties of statistical methods in this area.

5.5.1 An Unpatterned Covariance Model

We start with an unpatterned form proposed by Ford (1986) as a crude description of the covariance structure. Ford considered a model of the following form

$$\Sigma = \sigma^2 J + T \qquad (5.1)$$

where σ^2 represents between subject variation, J is the unit matrix $\underline{1}.\underline{1}^T$ and $T=\{\tau_{ij}\}$ represents the matrix of intra-subject variances & covariances.

In this model the τ_{ij}'s represent the important parameters of interest for studying inter-regional coupling. In studying data with this structure we would be trying to estimate correlations of the form

$$\theta_{ij} = \frac{\tau_{ij}}{(\tau_{ii}\,\tau_{jj})^{1/2}} .$$

The presence of a random subject effect σ^2 means that what we estimate from the raw data is

$$RHO_{ij} = \frac{\sigma^2 + \tau_{ij}}{\left((\sigma^2 + \tau_{ii})(\sigma^2 + \tau_{jj}) \right)^{1/2}}.$$

Thus, correlations based on raw data are influenced, not just by inter-regional association, but also by between subject variation.

Correlation analysis of adjusted data is an attempt to recover the form of T and hence the true θ_{ij} . It can be shown however, that this can present problems. In effect, important parameters remain confounded with nuisance parameters.

Using the following simple results (involving variables X_1 to X_4) concerning variances and covariances

$$C(X_1-X_2, X_3-X_4) = C(X_1,X_3) - C(X_1,X_4) - C(X_2,X_3) + C(X_2,X_4)$$

$$\& \qquad V(X_1-X_2) = V(X_1) + V(X_2) - 2*C(X_1,X_2),$$

simple formula can be derived for some forms of adjusted correlation.

For example, normalisation by subtraction of a subject mean of all regions in the data vector, results in correlations between regions i & j of the form

$$RHO_{ij} = \frac{\tau_{ij} - \tau_{i.} - \tau_{j.} + \tau_{..}}{\left((\tau_{ii} + \tau_{..} - 2\tau_{i.})(\tau_{jj} + \tau_{..} - 2\tau_{j.}) \right)^{1/2}}$$

where $\tau_{i.} = \Sigma_{j} \tau_{ij}/p$ and $\tau_{..} = \Sigma_{ij} \tau_{ij}/p^2$. Since a number of nuisance parameters are included in this formula, this demonstrates the confounding effect with covariance parameters of interest. Similar formula can be derived for other forms of adjustment. Ford (1986) derived a formula of this type for studying the partial correlation coefficients.

Some investigators (Worsley et al (1991), Horwitz (1984)) have suggested that a full unrestricted covariance matrix overparameterises the true correlation structure and that there is a simpler form for the intra-subject variances and covariance structure. Some of the models they propose are introduced in the proceeding sections.

5.5.2 Compound Symmetry and Sphericity

Compound symmetry represents one special form of covariance matrix. This can be expressed in the form

$$\Sigma = \sigma^2 I + \tau^2 J$$

where σ^2 is the between subject variance, I is the identity matrix, τ^2 is the within-subject variance-covariance parameter and J is the unit matrix.

This implies that the variance of all pairwise differences among variables, $Var(X_i-X_j)$ $(1\leq i<j\leq p)$, are equal. In addition, the repeated measurements are equicorrelated. A more general form is known as sphericity. This only requires that $Var(X_i-X_j)$=constant.

Since the sphericity assumption was rejected (see sections 4.1) we can presume that neither the compound symmetric or spherical models apply with these data and quite probably to imaging datasets in general.

5.5.3 Regional Clustering

A number of exploratory studies of the sample correlation structure have noted a degree of clustering among regions (Horwitz, 1984). This can be expressed in terms of a partition of the overall covariance matrix into C submatrices

$$\Sigma = \begin{vmatrix} \Sigma_{11} & 0 & \cdots & 0 \\ 0 & \Sigma_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Sigma_{CC} \end{vmatrix} .$$

In this case, variables within clusters will exhibit association, while variables in different clusters will be uncorrelated.

Three simple representations of this are illustrated in Figure 5.6. R_1 represents complete independence indicating 4 single region clusters. R_2 shows two clusters of two regions, one with a coupling coefficient of r_1, the other with r_2. R_3 shows one cluster all equally coupled. Such clustering type structures have been shown to be useful in devising simulation studies (MacCormack ,1991) and in studying the theoretical effect on partial correlations (Ford, 1986).

$$R_1 = \{\text{diag}(1,1,1,1)\} \qquad R_2 = \begin{matrix} 1 & r_1 & 0 & 0 \\ r_1 & 1 & 0 & 0 \\ 0 & 0 & 1 & r_2 \\ 0 & 0 & r_2 & 1 \end{matrix}$$

$$R_3 = \begin{matrix} 1 & r & r & r \\ r & 1 & r & r \\ r & r & 1 & r \\ r & r & r & 1 \end{matrix}$$

Fig 5.6 Representations of regional clustering

5.5.3 Linear Covariance Models

The final model which has been used in the context of imaging data is the general linear covariance model. Here we express the pattern of variances and covariances in a covariance matrix as a linear combination of other known matrices. That is

$$\Sigma = \sum_{i=1}^{t} a_i V_i$$

A simple case of this is the compound symmetric model of the previous section. Another is the spatial model described in section 5.4.

5.7 DISCUSSION

The study of interrelationships among regions is clearly very important if biological understanding of the brain is to advance. However, it is clear from some of the analysis in this chapter, that arriving at reliable biological conclusions from the results is not straightforward using these datasets. It seems that any effort to determine functional association between regions using correlation analysis must be tempered with sober realisation of the inherent confounding introduced into the analysis by prior adjustment of data. It has been suggested (Metter et al, 1984a) that such inferences can only be made from the study of multiple scanning vectors per subject, obtained under ideal conditions, before we can truly study intra-subject correlations. At the end of day, however, it should be remembered that reconstruction algorithms involve a degree of smoothing. This can clearly impose inter-pixel correlations locally and even globally which may undermine our aims.

Of all the simplified covariance structures described in section 5.6, the linear spatial model,

described in section 5.5, would appear most plausible. This is backed up by the results of the exploratory analysis of real data carried out in section 5.4. However, the adequacy of the three parameter model requires further study. From the data provided by Worsley we can investigate the adequacy of the model visually from a plot of $^1/_2\ V(X_i-X_j)$ against d_{ij} (Figure 5.7). It is apparent that there is still a great deal of unexplained variation around a best fitting line for these data. In practice other factors might be incorporated into the model. Indeed, we might replace the simple distance measures used by Worsley and with our SPECT data, by biological distance such as the lengths of neuronal pathways between regions.

A number of reasons could explain the presence of the spatial effect in imaging datasets. On the one hand, the variation might be genuinely due to underlying physiological reasons. Alternatively, it might in part be due to the measurement and/or reconstruction methodology used. It is entirely possible that the larger variance estimates for regional differences between distant regions over closer regions reflects some difficulty in matching the orientation through regions across a group of homogenous subjects - normal and diseased.

As a feature of imaging datasets, the spatial covariance form has important implications for the univariate analysis of regional means in within and between group problems - as illustrated in chapter 3,

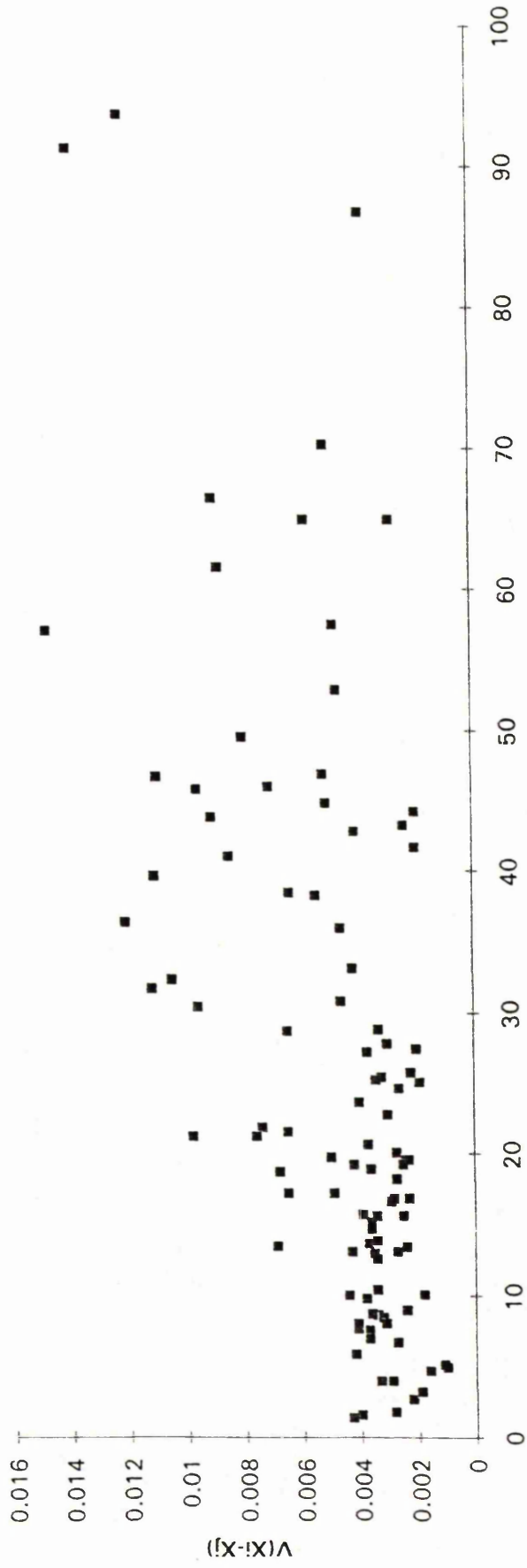


FIGURE 5.7 Worsley PET data: Plot of variance estimates of pairwise regional differences against distances between regions

especially, if the data is normalised to a single 'spared' region e.g. the occipital region. An implication of the spatial model is that, normalised regional variates (where the region is far from the normalising region) will have larger standard errors of mean than for closer regions. This will then make statistical significance between regions and between groups harder to detect. Similar problems arise with subject mean normalised data (Worsley, 1991).

CHAPTER 6

ANALYSIS OF LONGITUDINAL DATA

6.1 INTRODUCTION

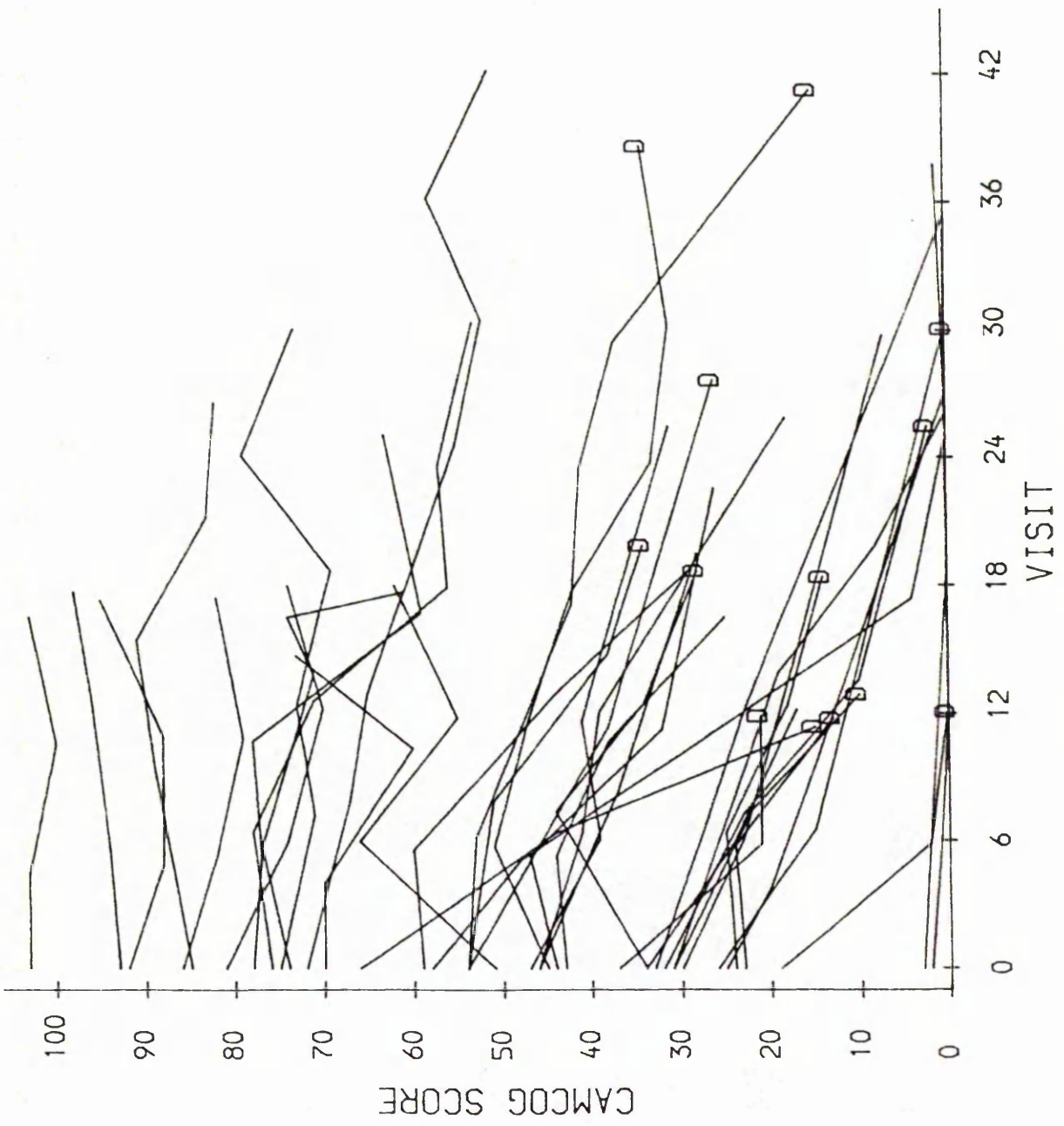
In DAT, the onset of the disease is usually unknown and difficult, if not impossible, to determine. Hence the length of time between onset and diagnosis is likewise difficult to determine. However, we do know that the illness progressively affects the individual in a number of different respects e.g. cognitive and behavioural function. This is illustrated figure 6.1 showing showing the decline in individual's CAMCOG (Roth et al, 1986) scores, a neuropsychologica measure of general cognitive function, over time from entry to the study.

Assuming some correlation with cerebral activity, we might expect to see a related change in the SPECT profiles of an individual over time. Only a few studies have looked into this problem (Barclay et al, 1984; Jagust et al, 1988).

6.2 A LONGITUDINAL DATASET

Regional mean/counts pixel data was observed in p=14 regions for n=17 subjects on t=4 different occasions - roughly six months apart. Table 6.1 contains regional subject means and standard deviations on each occasion. Just as between individuals, a feature of these data is the presence of large scalar difference between profile

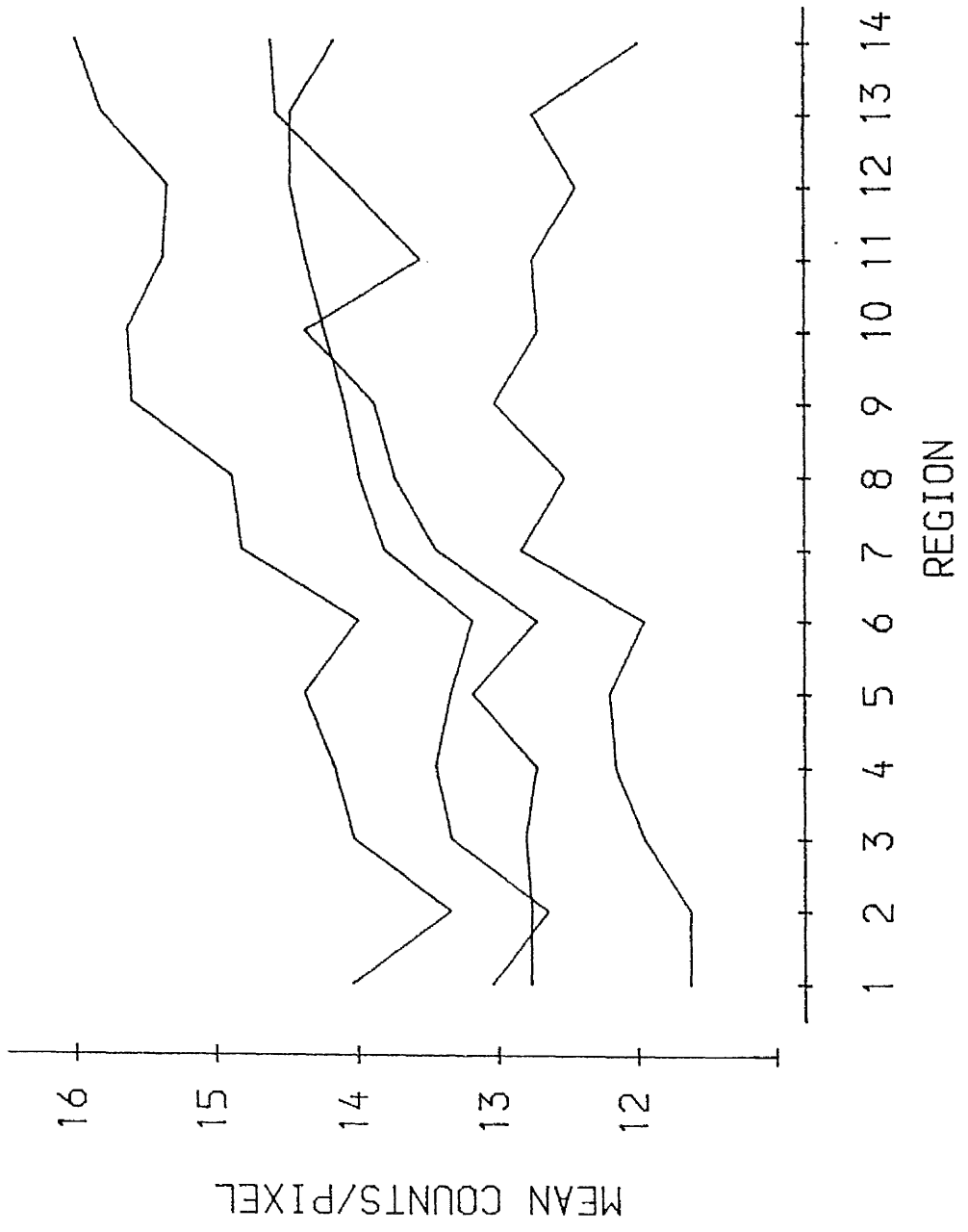
Figure 6.1 Plot of individual's CAMCOG scores against time in months from entry to the study.



REGION	TIME 1	2	3	4
RF	12.182	12.603	11.940	12.198
	2.037	1.432	1.441	1.051
LF	12.074	12.464	11.748	12.136
	2.043	1.504	1.320	1.153
RHF	11.993	12.390	11.838	12.073
	2.106	1.442	1.455	1.177
LHF	11.997	12.408	11.875	12.145
	2.166	1.555	1.435	1.306
RT	12.464	13.039	12.391	12.629
	2.281	1.495	1.296	1.213
LT	12.339	12.707	11.999	12.422
	2.073	1.498	1.329	1.222
RP	12.609	13.065	12.427	12.749
	1.972	1.349	1.219	1.141
LP	12.527	12.999	12.406	12.669
	1.974	1.383	1.217	1.282
RPT	12.850	13.453	12.773	13.197
	2.027	1.331	0.951	1.073
LPT	12.813	13.298	12.422	12.885
	2.047	1.320	1.143	1.288
RBG	13.444	14.168	13.215	13.439
	2.664	1.406	1.585	1.383
LBG	13.478	14.222	13.424	13.457
	2.508	1.315	1.314	1.334
RO	13.451	14.095	13.460	14.026
	2.082	1.188	1.095	1.382
LO	13.149	13.888	13.286	13.782
	1.902	1.173	1.100	1.324

Table 6.1 Regional means and standard deviations on each occasion.

Figure 6.2 Plot of ROI profiles on four occasions for one individual



SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F	TAIL PROB.	G-G PROB.	H-F PROB.	
Between Subject								
Mean	156139.34331	1	156139.34331		2401.15	0.0000		
Error	1040.42876	16	65.02680					
Within Subject								
Time	61.94305	3	20.64768		1.11	0.3561	0.3378	0.3422
Error	896.49283	48	18.67693					
Region	325.38302	6	54.23050		43.06	0.0000	0.0000	0.0000 *
Error	120.90357	96	1.25941					
Time x Region	7.12084	18	0.39560		1.97	0.0114	0.0796	0.0425 *
Error	57.87242	288	0.20095					
Hemis.	2.99601	1	2.99601		2.35	0.1451		
Error	20.43010	16	1.27688					
Time x Hemis.	0.05071	3	0.01690		0.11	0.9510	0.9320	0.9510
Error	7.06964	48	0.14728					
Region x Hemis.	3.66638	6	0.61106	2.67	0.0195	0.0528	0.0374 *	
Error	21.99874	96	0.22915					
Time x Region x Hemis.	1.24777	18	0.06932	1.33	0.1682	0.2396	0.1885	
Error	15.02124	288	0.05216					

Table 6.2 Repeated measures anova including time, region and hemisphere effects.
Asterisks denote significant effects at 5% level.

data on each of the different occasions. This is illustrated in figure 6.2.

6.3 A RM ANOVA APPROACH TO ANALYSIS

Both Barclay et al (1984) and Jagust et al (1988) analysed their data by normalising to the subject mean on each occasion and comparing occasions for each region separately using paired t-tests. For data on more than two occasions this approach could result in a large number of tests and the interpretation could complicated for reasons similar to those in section 3.6. As a first step we analysed our dataset using the univariate RM ANOVA, including time, region and hemisphere as within subject factors. The results of this analysis are summarised in Table 6.2. We can see from the results that there is a very significant region effect and some evidence of region x time and region x hemisphere interactions. However, in both these latter two tests there is some doubt, as the Greenhouse-Geisser adjusted p-values just fail to reach significance at the 5% level.

6.4 SELECTED CONTRASTS IN TIME

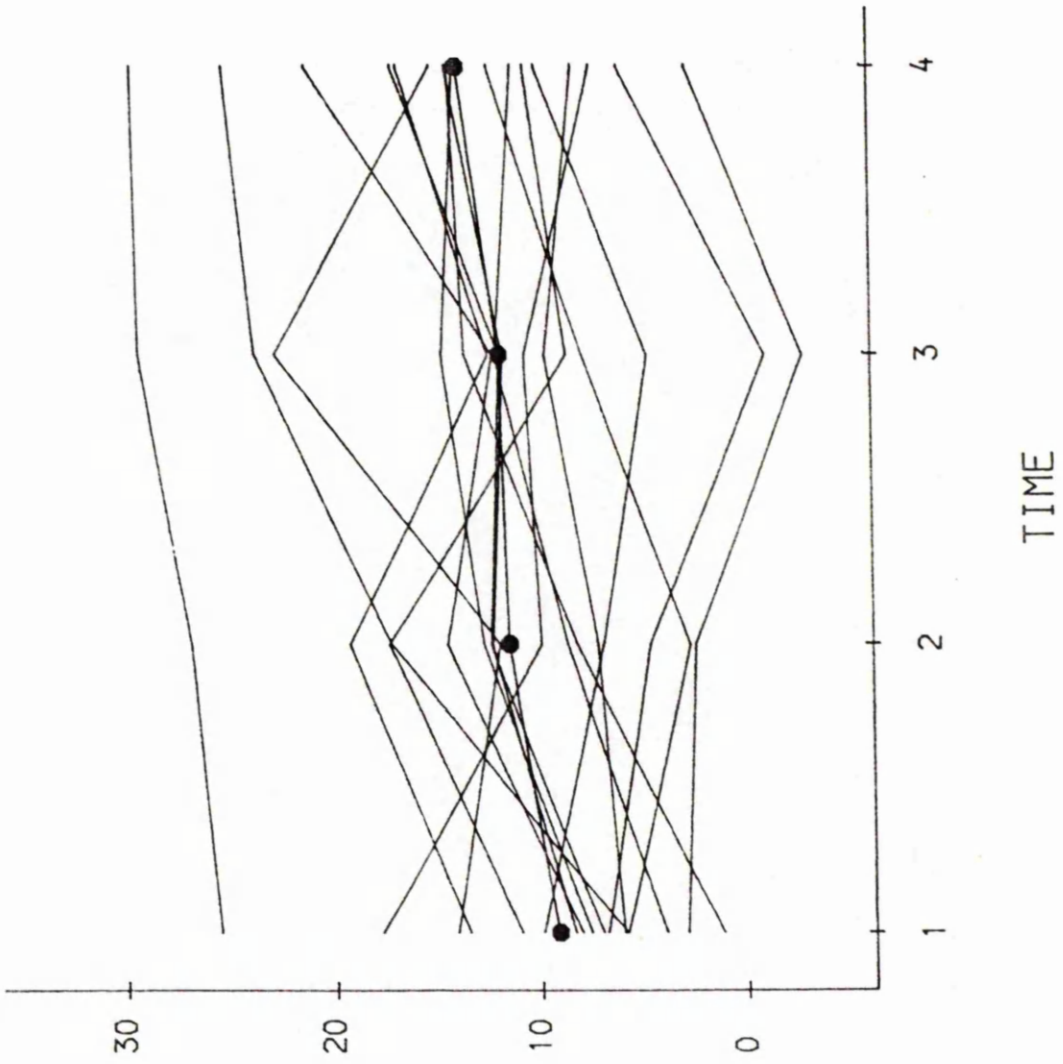
Assuming the region x time interaction is significant, we would naturally like to carry out a follow-up analysis of these data. In the previous analysis of the single scan data we noted that, for this disease group, the overriding difference from the normal

group was the contrast in activity between occipital and other regions. In order to see whether this feature was in any way correlated with time further analysis was carried out on the contrast

	RF	LF	RHF	LHF	RT	LT	RP	LP	RPT	LPT	RBG	LBG	RO	LO
Z^T =	(-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	0	0	+5	+5).

A plot of the individual contrast values over time are given in figure 6.4, together with the means at each time. There is considerable variability at each time point and clearly not all subjects conform to a stable or even obvious trend. Nonetheless, there is some indication of a positive trend in the mean values over time. A formal analysis of Z values was carried out using the rm anova approach - including time as the only factor. The results are given in Table 6.3. The overall test is very clearly significant, even after adjusting for violation of the sphericity assumption in time. The 3 d.f. time effect was further broken down into orthogonal linear, quadratic and cubic components using the 'ORTHOGONAL' option in P2V of BMDP. The test for linear trend was the only significant effect in this data. Thus there was evidence to suggest that on average the contrast between occipital and other regions decreases in roughly equal amounts over the time period (approximately 18 months) studied here.

Figure 6.4 Plot of Z contrast values against estimated time since onset of DAT.



SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F	TAIL PROB.	G-G PROB.	H-F PROB.	
Between subject								
Mean	9184.17507	1	9184.17507	64.41	0.0000			
Error	2281.60100	16	142.60006					
Within Subject								
Time	196.67504	3	65.55835	4.27	0.0094	0.0227	0.0176	*
Error	736.29971	48	15.33958					
Linear	185.63441	1	185.63441	6.96	0.0179	*		
Error	426.53734	16	26.65858					
Quadratic	0.41983	1	0.41983	0.05	0.8264			
Error	135.16371	16	8.44773					
Cubic	10.62081	1	10.62081	0.97	0.3385			
Error	174.59866	16	10.91242					

Table 6.3 Repeated measures anova of Z contrast in time with orthogonal decomposition into linear, quadratic and cubic trend. Asterisks denote significant effects at 5% level.

6.5 ADJUSTING FOR BASELINE STATE

In most cross-sectional studies we have little or no control over the sample of subjects who are recruited in each group. Thus we are likely to obtain a group of subjects who are at various stages of the disease. Now, it seems plausible that the latency period, between onset and diagnosis, could be an important factor in accounting for much of the between subject variability seen at time point 1 in figure 6.3. In cross-sectional studies an attempt is made to account for this, by categorising subjects in terms of the severity of the disease on the basis of presenting clinical and psychiatric symptoms.

The RM analysis of the previous section was augmented by inclusion of a categorical factor for the severity of the disease at the time of the first scan. In our dataset 3 cases were categorised as `minimal', 7 as `milds' and 7 as `moderate' in severity. The results are given in Table 6.4. Only the time effect is significant.

Alternativel we could some measure of the latency period (i.e. onset - diagnosis) of the disease as a covariate in relation to the vector of ROI responses. In the Alzheimer study, an estimate of time since onset was obtained from a relative, close friend or someone who had been in contact with the subject over a period of time and had notionally observed the decline in their condition. We plotted the Z values at time 1 against this data in (Figure 6.4). There is no visual evidence to suggest that the relationship is anything but weak. There

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F	TAIL PROB.	G-G PROB.	H-F PROB.	
Between subject								
Mean	6982.90374	1	982.90374	47.22	0.0000			
Severity	211.39522	2	105.69761	0.71	0.5063			
Error	2070.20578	14	147.87184					
Within Subject								
Time	197.76442	3	65.92147	4.08	0.0125	0.0306	0.0189	*
Severity x Time	57.49051	6	9.58175	0.59	0.7342	0.6617	0.7045	
Error	678.80920	42	16.16212					

Table 6.4 Repeated measures anova of Z contrast in time and by severity grouping. Asterisks denote effects significant at 5% level.

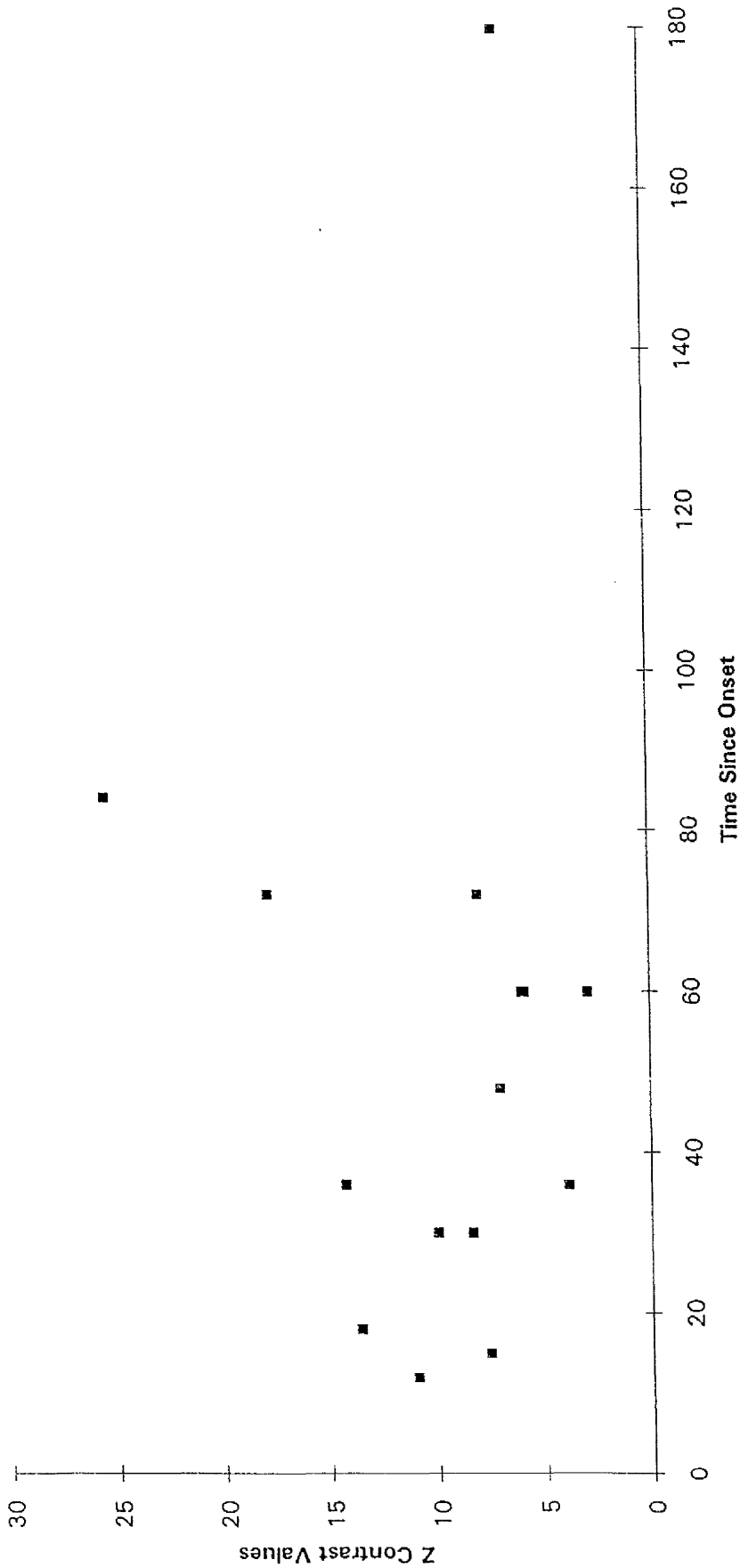


FIGURE 6.5 Plot of Z contrast values against time since onset of DAT

are obvious problems with the onset data. The estimates are highly subjective since they are prone to differing degrees of accuracy among individuals. The data point on the extreme right of figure 6.4 seems unrealistically high.

6.6 DISCUSSION

Compared to cross-sectional analysis, the analysis of serial data presents a set of additional statistical problems.

In situations where we have different numbers of scans/subject, the RM ANOVA approach may not be applicable since we would be dealing with unbalanced numbers of repeated measurements. The antedependance approach of Kenward (1987) is useful in the single variable problem and might be of some use here also. If scans at each time point were reasonably strongly corrlated, a similar sort of analysis could be incorporated into the framework of the RM ANOVA by taking successive differences of the form $\underline{X}_{j+1} - \underline{X}_j$. If no change had occurred in an individual or a group of individuals, the individual or group mean difference would be constant across regions. This approach might also be used as a way of adjusting for basline differences.

CHAPTER 7

CORRELATING DIFFERENT DATASETS

Many disciplines are involved in brain research which can result in the collection of a vast array of different types of data on any one individual. One of the concepts of using a multidisciplined team in the Alzheimer study was that we could study the disease both in-vivo and in-vitro and the correlation between data at distinct levels of biological resolution for the same subjects.

In this chapter we will look at some examples of the approaches taken to correlate data collected in the study from the different disciplines . We will look at the correlation of ROI SPECT data with neuropsychological test scores and with senile plaque count data. In each of these problems we essentially are interested in comparing two multivariate datasets; a blood flow dataset X (n x p) and a neuropsychological or plaque density dataset Y (n x q), with p and q variables in each, observed in i=1,...n common individuals.

7.1 SPECT AND NEUROPSYCHOLOGICAL TEST SCORES

Four neuropsychological test scores were chosen from the Alzheimer database for this analysis and are described in Table 7.1. Only Alzheimer subjects were looked at in this analysis.

VARIABLE	DESCRIPTION
Y_1	Measure of Praxis from CAMCOG.
Y_2	Graded Naming Test : Total errors on test.
Y_3	Verbal Fluency: composite of 6 subscores.
Y_4	Memory: composite score of visual/pictorial memory.

TABLE 7.1: Neuropsychological Test Variables

One simple approach to study these data is to calculate correlation coefficients between each regional variable and each neuropsychological variables (Montaldi, 1991). Because of the subject effect, this only makes sense if the SPECT data is adjusted prior to correlating the data. This can be done by normalisation as described in chapter 3. For this analysis the regional data had the subject region mean subtracted. Table 7.2 displays individual sample correlations between each normalised regional variable and each of the neuropsycholgical test scores in table 7.1. Initial scrutiny of the size of the correlation coefficients suggests that any association that exists is weak. Formal analysis the correlations was carried out by testing the hypothesis of zero correlation. Because of the multiplicity of testing involved, individual significance was assessed using Bonferroni adjusted alpha levels. A number of statistically significant correlations were identified.

As in the analysis of means, it is important to remember that other forms of normalisation will likely

	Y_1	Y_2	Y_3	Y_4
RF	-0.06	0.16	0.12	0.04
LF	-0.01	0.14	0.13	0.06
RHF	-0.02	0.03	0.14	0.06
LHF	-0.04	0.14	0.02	-0.02
RT	0.15	-0.09	0.04	0.24*
LT	0.14	0.27*	0.02	0.12
RP	0.19	-0.18	-0.02	0.23*
LP	-0.01	-0.22*	-0.09	-0.01
RPT	0.23*	-0.01	0.01	0.04
LPT	0.17	-0.03	0.01	-0.14
RO	-0.31*	-0.12	-0.14	-0.23*
LO	-0.28*	-0.09	-0.12	-0.18

Table 7.2 Correlation coefficients between adjusted (to subject mean) imaging data and neuropsychological test scores. Asterisks denote significant correlations at the 5/48% significance level.

give different results and therefore that interpretation of any results must be in terms of normalised data; not in terms of the original data.

We looked at other methods of correlating these data. Regression analysis was carried out relating each neuropsychological variable to the blood flow profile data. In the interest of simplifying the amount of

analysis involved, we first reduced dimensions by first transforming to regional hemispheric sums and differences. In the analysis below hemispheric sum variables are prefixed by 'S' e.g. SO=occipital sum, and difference variables by 'D' e.g. DO=occipital difference. Multiple correlation coefficients between each neuropsychology variable and the set of regional sums and differences separately are shown in Table 7.3 together with a test of significance. In contrast to the previous analysis the hemispheric sum data were not adjusted prior to analysis, as the regression analysis accounted for the scalar effect in the interrelationships among region variables.

Neuropsych -ological Variable	SUMS R^2	SUMS Sig. Level	DIFFERENCES R^2	DIFFERENCES Sig. Level
Y_1	.26	.001	.07	.526
Y_2	.09	.378	.20	.015
Y_3	.04	.851	.05	.708
Y_4	.13	.161	.09	.366

Table 7.3 Multiple correlation coefficients and p-values between each neuropsychological variable and the vectors of right/left sums and differences.

(a)

Regression Coefficient	Parameter Estimate	Standard Error	T-Ratio	P-value
Constant	4.487	2.983	1.63	0.109
SF	0.491	0.583	0.84	0.403
SHF	-1.022	0.622	-1.64	0.105
ST	0.522	0.458	1.14	0.258
SP	0.412	0.557	0.74	0.463
SPT	1.079	0.426	2.53	0.019
SO	-1.342	0.366	-3.67	0.000

(b)

Regression Coefficient	Parameter Estimate	Standard Error	T-Ratio	P-value
Constant	8.588	0.418	20.56	0.000
DF	1.568	1.657	0.95	0.347
DHF	-2.338	1.408	-1.66	0.101
DT	-3.579	0.971	-3.69	0.000
DP	1.376	1.082	1.27	0.208
DPT	-0.035	0.949	-0.04	0.971
DO	0.181	0.888	0.20	0.839

Table 7.4a, b Regression parameters for (a) praxis with hemispheric sums and (b) graded naming test with hemispheric differences.

There is evidence from this analysis that the praxis variable is significantly (p=0.001) related to some or all of the hemispheric sums and that the graded naming scores are related (p=0.015) to some or all of the hemispheric differences. The regression model relating the praxis variable to hemispheric sums is given Table 7.4a and graded naming to hemispheric differences in Table 7.4b. From inspection of the t-ratios and p-values of each coefficient in these models, it would appear that only some of the variables were related to the neuropsychological variables. Further analysis of these data using a stepwise variable selection gave the following reduced model for praxis

$$Y_1 = \underset{(2.89)}{4.44} + \underset{(.294)}{1.287} \times SPT - \underset{(.317)}{1.121} \times SO$$

which had an $R^2=0.225$, and for graded naming

$$Y_2 = \underset{(.380)}{8.87} - \underset{(.904)}{3.14} \times DT$$

which had an $R^2=0.142$.

The analysis would tend to suggest that praxis scores were related to a contrast between posterior-temporal and occipital regions, while graded naming scores were inversely related to the degree of asymmetry in the temporal region.

An extension of this analysis was carried out using canonical correlation analysis, relating all neuropsychological variables to all imaging variables. The first canonical variate between hemispheric sums & differences together and neuropsychological test scores is highly significant (p=0.0106). The same analysis repeated on sums and differences separately showed that sums were highly significant on the first canonical variable (p=0.0064) while differences were not (p=0.3610). This latter result is not totally consistent with Table 7.3. This may be accounted for by the fact that 5% significance levels were used in that analysis and that if Bonferroni significance levels were used in interpreting the results of table 7.2, the two would be more consistent.

In order to understand the nature of the relationship between these two datasets better, the canonical variate coefficients and the correlation of the canonical scores to original variables were studied (Table 7.5). From the coefficients it can be seen that the canonical variate for the neuropsychology variables is dominated by praxis (Y_1). The signs of the canonical variate coefficients for the bood flow data would indicate that this describes the contrast between occipital and posterior-temporal regions. This is similar to the results obtained from the multiple regression analysis taking the neuropsychological data one variable at a time.

Neuropsych			SPECT		
Variable	Can.Coef.	Corr.	Can.Coef.	Corr.	Variable
Y_1	1.294	(.92)	0.554	(.60)	Fr
Y_2	0.058	(.08)	-1.225	(.60)	HF
Y_3	-0.188	(.19)	0.480	(.66)	T
Y_4	-0.446	(.36)	0.660	(.67)	P
			2.223	(.74)	PT
			-2.097	(.47)	Oc

TABLE 7.5 Coefficients of canonical variates and (correlation coefficients) between canonical variables and original right/left sum variables.

There is an inherent difficulty in interpreting the canonical coefficients in this way because the imaging variables are themselves so highly correlated. However, in a rerun of the analysis, excluding posterior-temporal and occipital regions, no relationship between the two groups of variables could be found ($p=0.1181$). This would tend to back up the conclusions from analysis on the full set of regions

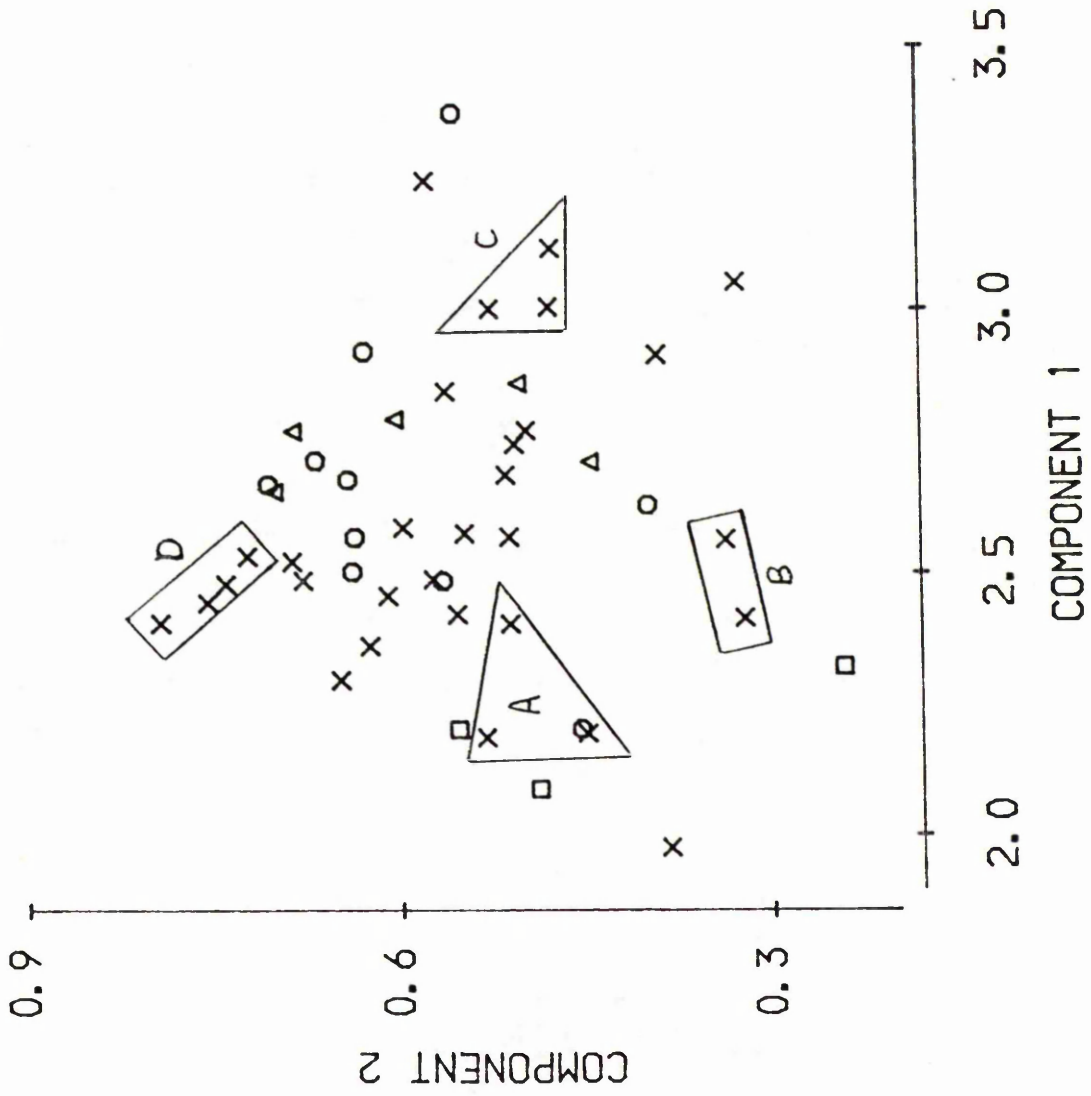
7.2 AN APPROACH BASED ON EXTREME SUB-GROUPS OF SUBJECTS

A more heuristic approach to the problem considered

above was investigated in the Alzheimer study and is reported below. This involved identifying cases with extreme SPECT profiles and comparing these cases in terms of corresponding clinical and neuropsychological data. SPECT and neuropsychological data from 48 DAT cases were used in this investigation.

Variation in regional SPECT patterns was investigated using PCA based on the covariance matrix of normalised ROI data. The data was normalised to occipital activity prior to the analysis. A two dimensional summary of inter-subject variation was obtained from the PCA using the first two principal components. These accounted for 70.4% and 12.2% (total 82.6%) of the variation. The coefficients for these components are given in Table 7.6, arranged according to component 2. For component 1 the coefficients are of the same sign and roughly the same magnitude. Thus, this describes an overall regional average measure of relative activity. The second component corresponds to a contrast between a posterior-temporal/parietal grouping and a frontal/higher-frontal grouping of regions. PC scores are plotted in figure 7.1, with different symbols used to identify clinical severity. The moderate group, being the largest of these groups, tends to dominate. Nonetheless, it is perhaps possible to detect clustering of subjects in the minimal and severe groups on component 1 although it is clear that the numbers involved and degree of the differences are small.

Figure 7.1 Plot of scores on first two principal conponents with four sub-groups highlighted. Clinical severity superimposed: Δ-'minimal', O- 'mild', X- 'moderate', □- 'severe'.



Number	Regions	Component 1	2
1	Posterior Temporal (Right)	.17	.50
2	Posterior Temporal (Left)	.24	.51
3	Parietal (Right)	.28	.29
4	Parietal (Left)	.28	.38
5	Frontal (Right)	.31	-.26
6	Frontal (Left)	.32	-.24
7	Higher Frontal (Right)	.38	-.26
8	Higher Frontal (Left)	.36	-.22
9	Temporal (Right)	.38	-.09
10	Temporal (Left)	.38	-.06
% VARIABILITY EXPLAINED		70.4	12.2

TABLE 7.6 Coefficients from principal components analysis

On the basis of figure 7.1, four clusters of cases from the moderate group were selected for further study. These were chosen from extreme quadrants in the north, south, east and west areas and are labelled as A, B, C or D in figure 7.1. The individual unnormalised ROI profiles are displayed for these subjects in figure 7.2 a-d. On visual evidence these show broad similarity within sub-groups and dissimilarity between sub-groups. Images for

Figure 7.2 Profile plots for each of the subjects in each sub-group.

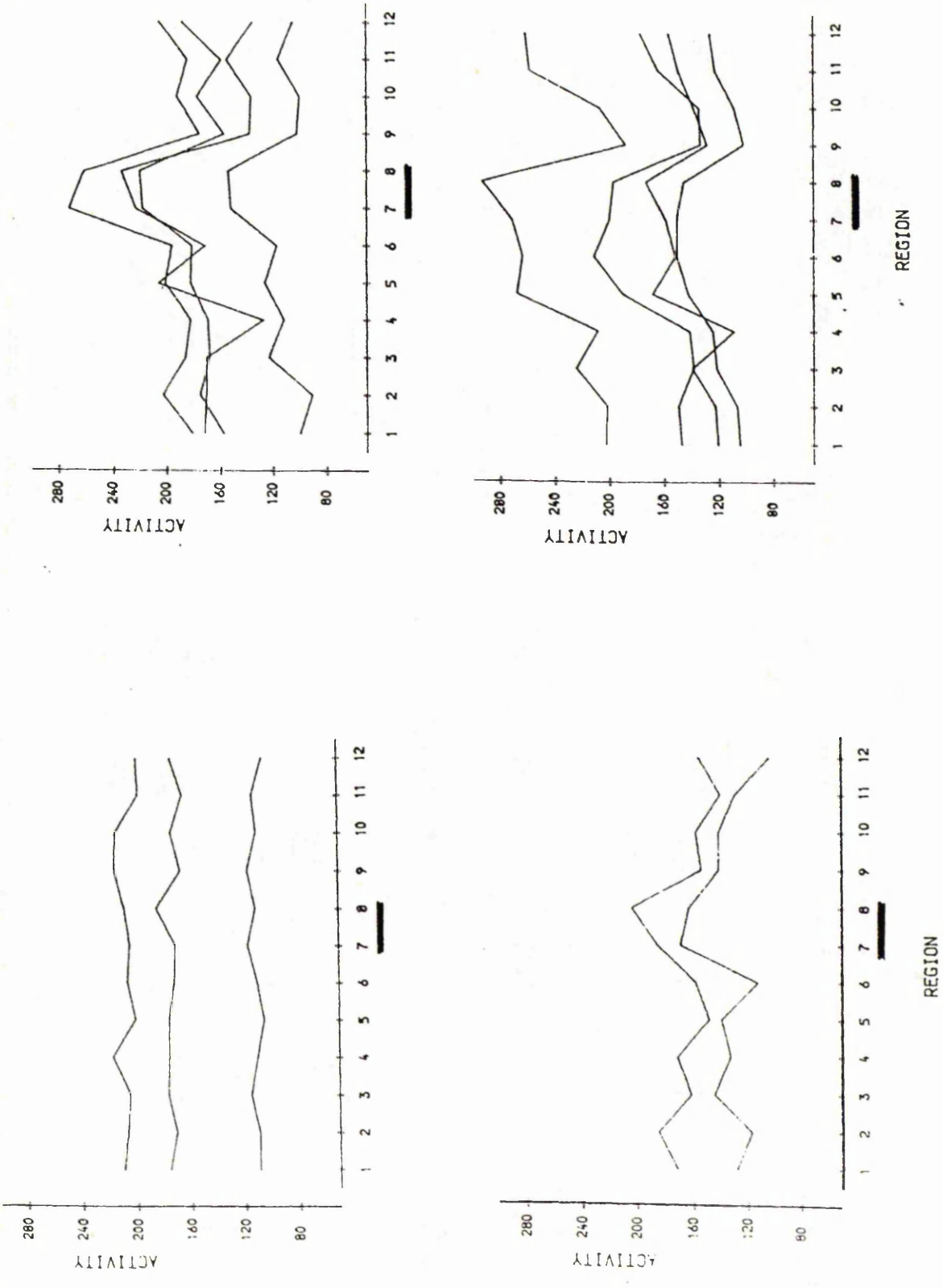
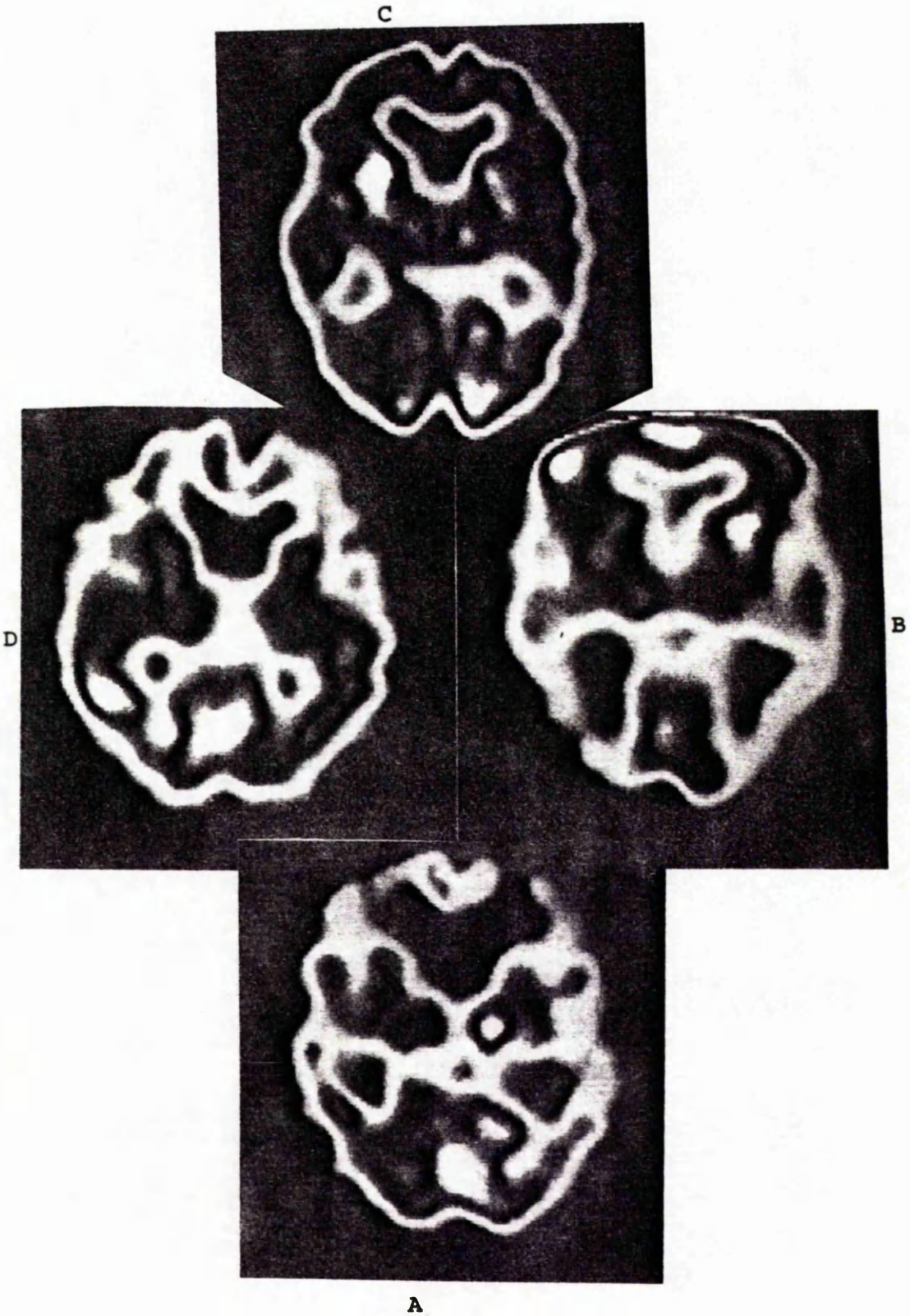


Figure 7.3 Images for one subject in each of the sub-groups of figure 7.1.



one subject in each sub-group were selected for comparison. These are shown in the montage in Figure 7.3a-d. From a clinical point of view there is little doubt surrounding the distinctions between them.

SUB-GROUP	CASEID	AGE (years)	BLOCK DESIGN	DIGIT SPAN	WORD FLUENCY	MMSE
	1	76	0*	5	2.0	5
A	2	84	0	5	1.5	*
	3	61	0*	0*	0*	1
B	1	68	0*	0*	0*	1
	2	75	0	4	6.0	3
	1	77	0	5	3.5	6
C	2	85	0*	0*	0*	8
	3	77	0	5	3.5	6
	1	79	0	5	3.5	12
D	2	91	0*	5	2.0	7
	3	89	0	4	6.5	14
	4	51	8	5	0*	18

TABLE 7.7 Neuropsychological test data for individual cases in each sub-group. '*' indicates missing data or test not completed. Zero imputed as most likely value on test.

The analysis proceeded by comparing the neuropsychological profiles of the four sub-groups. The neuropsychological variables chosen for this purpose

included an index of overall cognitive function (the mini mental state examination (MMSE)) and three other variables reflecting more focally organised cognitive functions (verbal fluency, digit span and block design). Age was also considered. These data are given in Table 7.7. From the range of the MMSE scores, we see that all patients were in the moderate to severely demented range of the index. Verbal fluency and Block Design, chosen to reflect predominantly anterior and predominantly posterior functioning (as indicated by component 2), showed no obvious differences between the four sub-groups. On a measure of immediate memory (Digit Span) there were no obvious differences between the four groups. On all tests subjects scored low. Age did not appear to explain any differences between the sub-groups either.

Consequently, there is little evidence, on neuropsychological grounds, to distinguish between these four sub-groups. Thus, there little indication, from this analysis, to suggest correlation between SPECT data and neuropsychological data.

These results initially surprised the radiologists involved in the study since the differences between the images in the four groups would normally be associated with clinical differences. One of the reasons for this is that we studied a group of DAT subjects who were clinically advanced in terms of severity. It has been noted that neuropsychological tests are not always sensitive enough to study the cognitive profiles in

subjects at advance stages of a disease; the scores often exhibiting floor effects - as is the case above. Another reason is that differences between the images reflect heterogeneous development of the disease which can no longer be linked to simple neuropsychological data. A third reason is a purely practical one. In figure 7.3, it is worth noting that the ventricles differ markedly in size from image to image. It is possible that the slices and orientation of the slices, from which the ROI data was extracted, did not strictly match for all subjects. This is a measurement problem using these techniques and is discussed further in chapter 8.

7.3 SPECT Counts v's Plaque Counts

The previous analysis dealt with the comparison of two different in-vivo datasets. We will now compare in-vivo SPECT counts with in-vitro cell degeneration, as measured by plaque count density. Plaque data for twelve subjects, with clinical Alzheimer's Disease, who had also had SPECT images carried out close to death. Plaque data was obtained from right and left frontal, temporal and occipital regions corresponding to the same regions imaged with SPECT. We thus had 6 anatomically matched SPECT and plaque data values for each subject for analysis.

In comparing SPECT to neuropsychological test scores we carried out a between subject analysis. The main issue for this comparison was whether to adjust the SPECT

and/or plaque data, as discussed previously, prior to analysis. There is some evidence of scalar differences in the plaque data similar, although to a lesser extent, to SPECT data. One feature of interest in relating these two datasets is the relationship between right and left measurements. Sample correlations are given in Table 7.8. These are all quite small, and on testing the null hypothesis of zero correlation, none were found to be significant.

		SPECT		
		Front	Sup. Temp	Occ.
PLAQUE	Front	.37	-.03	.32
	Sup.Temp.	.20	.19	.21
	Occ.	-.13	.06	-.11

TABLE 7.8 Correlation coefficients of right/left ratios. Ratios analysed on the Log scale.

Thus. there was insufficient evidence available to suggest any quantitative and hence biological link between SPECT and plaque data. With this in mind it is interesting to note an alternative approach to the analysis of datasets of this type (Dewar et al, 1991,

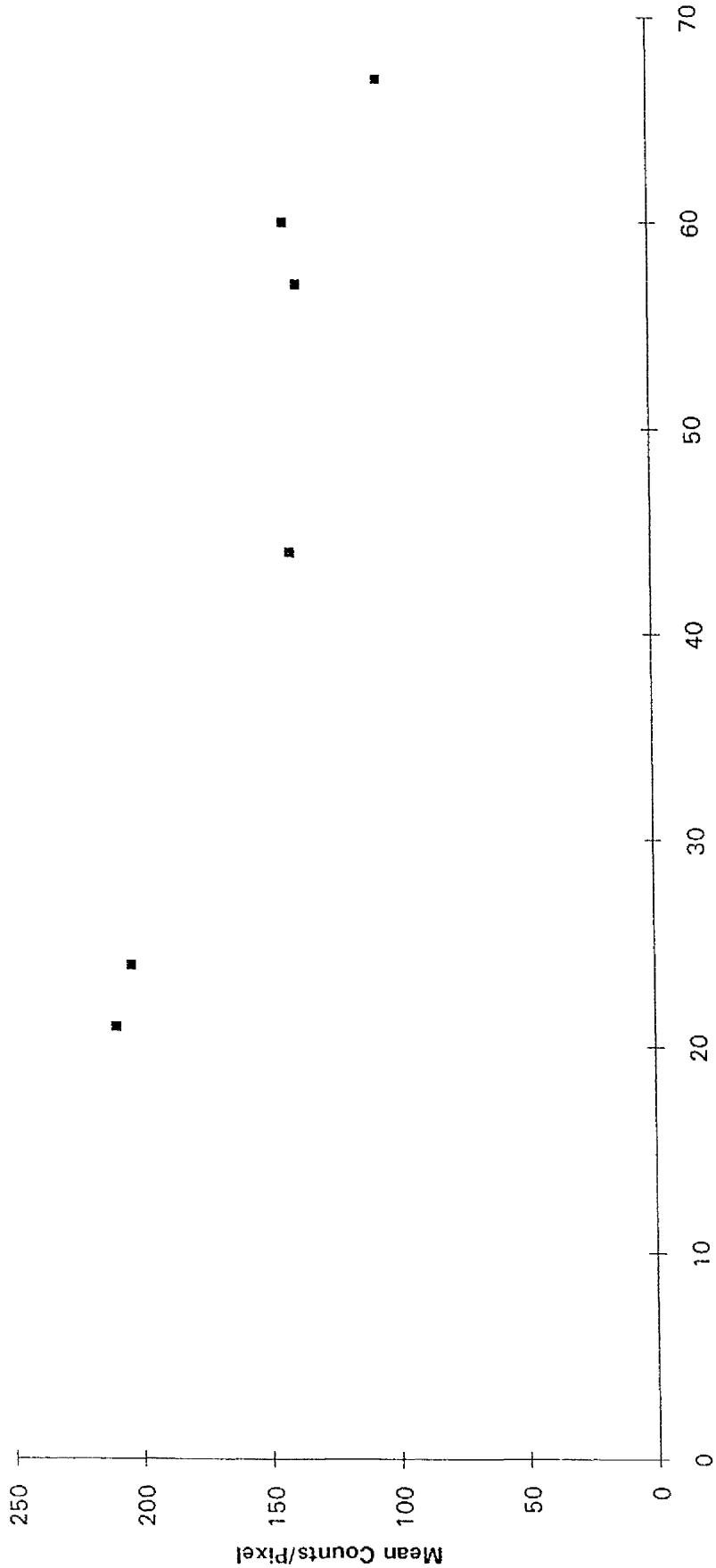


FIGURE 7.4 - Single Individual's Data

SUBJECT	CORRELATION	P-VALUE
1	-.96	.003
2	-.92	.010
3	-.78	.069
4	-.76	.079
5	.54	.267
6	.54	.269
7	-.50	.316
8	-.39	.441
9	-.26	.623
10	-.19	.723
11	-.12	.823
12	-.05	.925

TABLE 7.9 Within-subject correlations for each individual.

Duara et al, 1991). This involves the study of within-subject correlations; relating regional imaging data with plaque data for each subject separately across regions. We will consider the merits of this approach by applying this analysis to the data above.

For one of the subjects in our dataset plaque counts were plotted against SPECT count data figure 7.4. On the face of it, the two variables appeared to correlate quite

strongly (r=-0.96), with high ROI corresponding to low plaque numbers and vice-versa. Table 7.9 provides similar correlation estimates for the rest of the subjects. These are given in descending order of magnitude. The initial impression is that, for many of the subjects, the two variables appear strongly correlated, while for others the association is either very weak or negligible. In addition, in the majority of cases the correlations are negative - as in the example above.

In Dewar et al. (1991) and Duara et al. (1991), a t-test was used to assess the statistical significance of the observed correlations. This has been done for these data and presented in table 7.9. In both cases we would conclude that only two cases show statistically significant association at the nominal 5% level.

A number of aspects make this approach highly dubious. For one thing, the regional data (for both SPECT and plaques) will probably be themselves correlated. They cannot be considered as having arisen from a random sampling scheme, as assumed in correlation analysis. In addition, the regional data are structured measurements analagous with fixed levels of a factor (Region). On this fact alone, it is likely that different signs and magnitudes of `correlations' could be found simply by choosing different regions for the analysis.

7.4 DISCUSSION

Apart from statistical considerations in these

analyses there are als some important biological considerations. In certain neurological conditions, metabolic or blood flow deficits in one area may manifest themselves in a completely different area. An example is cross cerebellar diaschisis (Ell, 1990). Although more commonly associated with stroke, this fact is worth bearing in mind before placing any biological interpretation on the results of this analysis.

CHAPTER 8

OTHER TOPICS OF INTEREST

8.1 DISCRIMINATION PROBLEMS

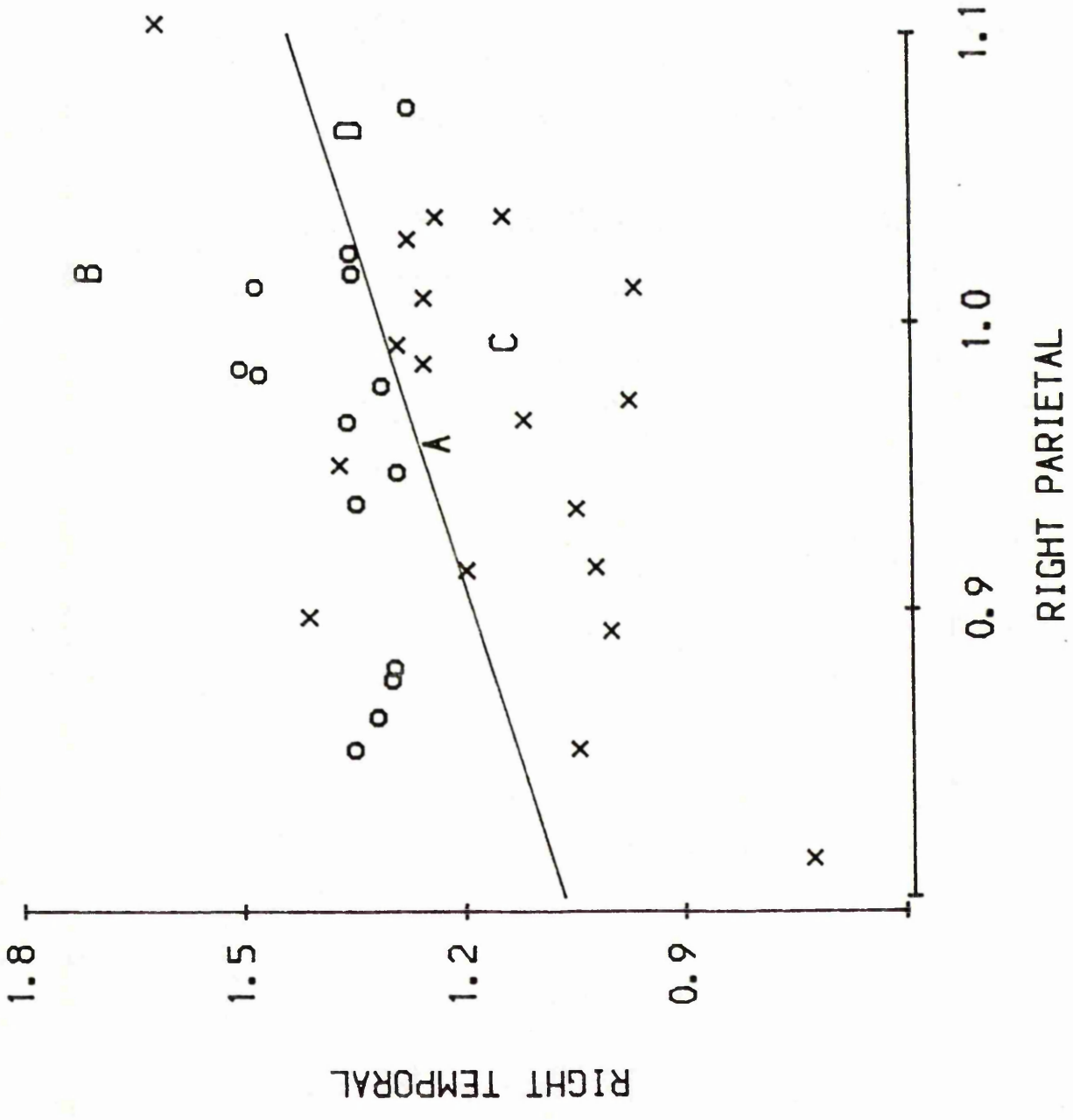
An important and practical application of imaging techniques is in problems of discrimination where information extracted from images might be used for diagnosis e.g. between normal and diseased states, or for screening individuals for inclusion in an experimental study or for deciding on suitable courses of prognostic action e.g. surgery. CT and MRI imaging are already used in some diagnostic and prognostic applications where qualitative features in the image e.g. infarcts, can be identified. To date the use of functional inaging in these problems has been mostly studied in terms of qualitative features (Gemmell et al, 1987; Testa et al, 1988; Jagust et al 1987; Smith et al, 1988), although some work has looked at the use of quantitative data (Clark et al, 1991; Ford et al, 1991).

We have investigated the use of SPECT data to address some of these problems in the Alzheimer study. Some of this work has been presentated as talks or posters at neuroscientific and general medical conferences and is presented below.

8.1.1 Discriminating Between Normal and Diseased Groups

McCrory and Ford (1990) studied the use of linear

Figure 8.1 Plot of normalised temporal against normalised parietal activity. X - Normal', O - DAT subjects.



discriminant analysis in a SPECT dataset consisting of twenty normal and sixteen early dementing DAT subjects. A variable selection routine was applied to these data as a first step. Normalised right parietal and right temporal regions were selected as providing best best discrimination between these groups. A plot of the data for these two regions is shown in Figure 8.1 with the linear disccriminant rule superimposed . Four additional cases were added to the plot for use as a training set. These are labelled A-D in figure 8.1. On the basis of visual inspection alone these data suggests that there is reasonable separation between the groups.

		True Group	
		Normal	**DAT**
Predicted Group	**Normal**	15	1
	DAT	3	13
	% Correct	83.3%	92.9%

Table 8.1 Forced classification table.

For classification purposes, a common approach is to assign cases to one of the design groups on the basis of the size of the estimated posterior probability. Using this approach, the misclassification rates for this data

are given in table 8.1. As expected from figure 8.1, the overall success of the classification rule for the desgn set is high, at 87.5%; with only one normal case being incorrectly assigned to the DAT group. For the training data, cases A, C & D would be assigned to the normal group while case B would be assigned to the DAT group.

In this example we assumed equal covariance matrices and prior probabilities. This latter assumption is probably given the incidence rates in table 1.2.

8.1.2 Uncertainty in assigning individuals to particular groups

A feature of this approach is that the predicted group membership involves forced classification to one or other of the groups, often using a value of 0.5 for the posterior probability as cut-off value.

There are a number of reasons why this is not a totally satisfactory approach for these data. In the first instance the sampling variability associated with the data means that there is some uncertainty about the true posterior probability for each individual. As we will see in the next section the ROI approach to data extraction suffers from inherent sources of measurement error. This is particularly important to remember in situations where an estimate is close to 0.5. Finally, it should be remembered that the diagnosis of DAT using clinical data is not totally reliable; only post-mortem data can confirm in-vivo diagnosis. Consequently, it is

entirely possible that some cases may be wrongly included in the design dataset. This is a problem of an `incorrectly classified design set' (Hand, 1981).

Recent research on quantifying the uncertainty in such situations includes the work of Schaafsma & Van Vark (1979) and Ford & Critchley (1982) and Rigby (1982). This involves constructing interval estimates for the posterior-probabilities (Rigby, 1982; Critchley and Ford, 1985; Critchley et al, 1988), for each indiviudal, which can be taken into account in the classification decisions.

Subject	**Pr.(Normal|Data)**	**Interval Estimate**
A	.59	(.46 , .73)
B	.02	(.01 , .41)
C	.86	(.66 , .97)
D	.69	(.36 , .90)

Table 8.2 Posterior probabilities and interval estimates

We used Rigby's approach to construct interval estimates of the posterior probabilities for the cases in our training set. These are given in table 8.2. As mentioned already, the forced classification approach would mean that cases A, C and D were classified as normal while B would be classified into the DAT group.

However inspection of the interval estimates suggests that the decison, in some cases, it is not so clear cut. In particular, the interval estimates for cases A and D also include 0.5. Study of other medical information would be adviseable in these cases.

Although cases B and C seem more straightforward no decision would be wisely taken on the basis of the imaging data alone. Even so, these results suggest that functional data may well be a useful addition to other medical information when takinhg diagnostic decisions.

8.1.3 Discriminating between DAT and Other Dementias

As mentioned in chapter one, clinical diagnosis of different dementias can often be quite difficult. In some situations, part of the diagnosis may require short to long term monitoring of an individual before different dementias can be distinguished. Given the results of the previous section, it is interesting to investigate whether differences can be detected between different dementias using their blood flow data. For this purpose we will compare our group of N=79 DAT subjects and a group of N=5 subjects where the dementia characteristics are secondary to other symptoms e.g infarcts. This latter group corresponds to the secondary DAT group in the Alzheimer project. Unfortunately the small number of subjects in the secondary dementia group prohibits a full discriminant type analysis as section 8.1.1. However, a cursory analysis of this question may be carried using

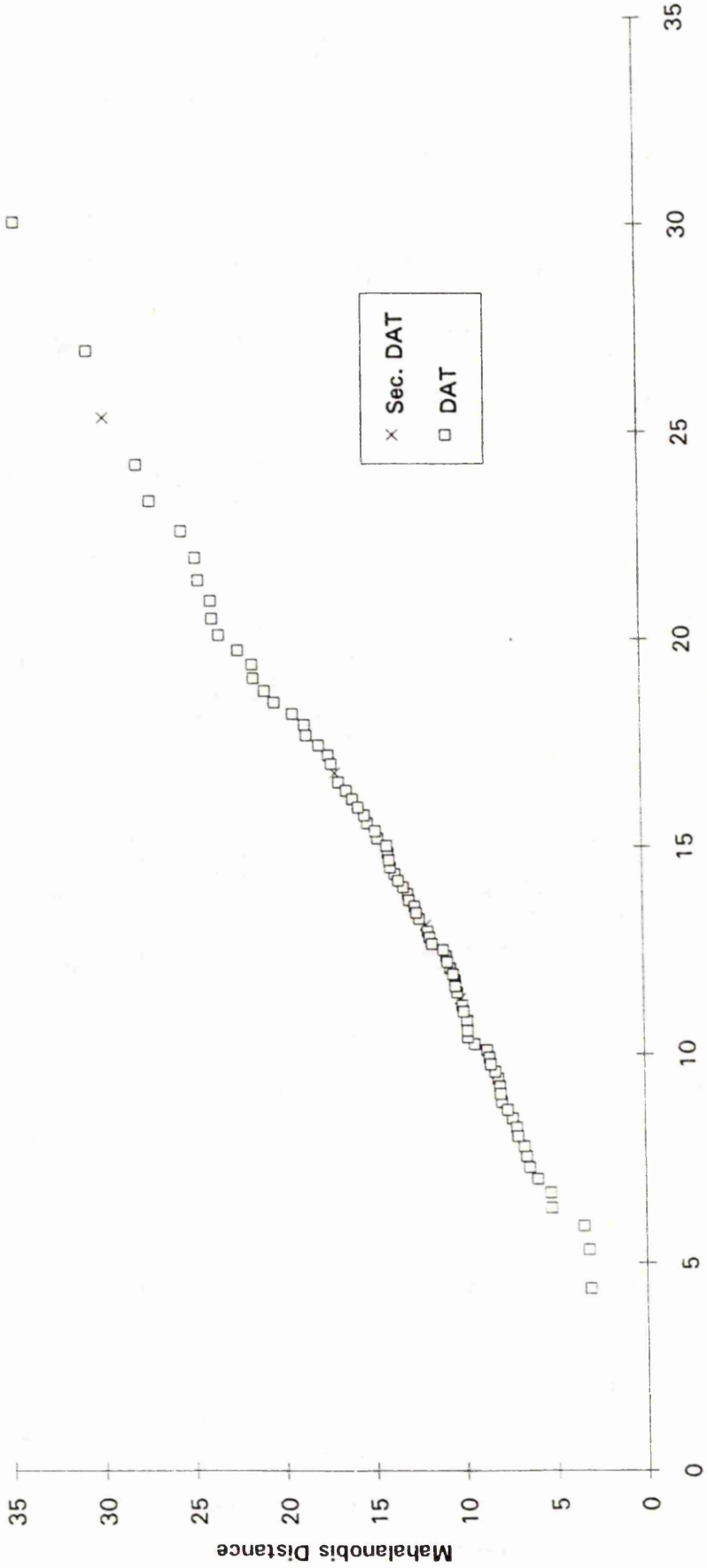


FIGURE 8.2 Probability-plot to compare DAT and Secondary DAT subjects

the fact that discrimination is just a means of analysing describing the multidimensional scatter of data between the groups. A useful measure of the scatter of data in multidimensional space is in terms of Mahalanobis distance. Distances, from the DAT group mean vector, were calculated for all subjects and displayed in a probability plot (figure 8.2).

Almost all of the secondary cases (labelled ?) lie in or around the main body of the DAT cases (labelled ?). This tells us, at the very least, that there is evidence of greater scatter around the average DAT profile in the DAT group than the secondary group and, at the most, that the secodary group is very much like the DAT group. However, this is the extent of the information that can be taken from this analysis.

8.1.4 Screeening Individuals

A related problem to those above is that of screening individual's functional data for suspicious features which might preclude their inclusion in a brain study; as with CT in the Alzheimer study. An important assumption with functional data in this situation is that standard reference ranges can be identified, within which the large majority of normal profiles would be located. Any case not lying in reasonable proximity to this range might therefore be considered as a potentially non-normal case, although this would probably have to be judged in combination with other data.

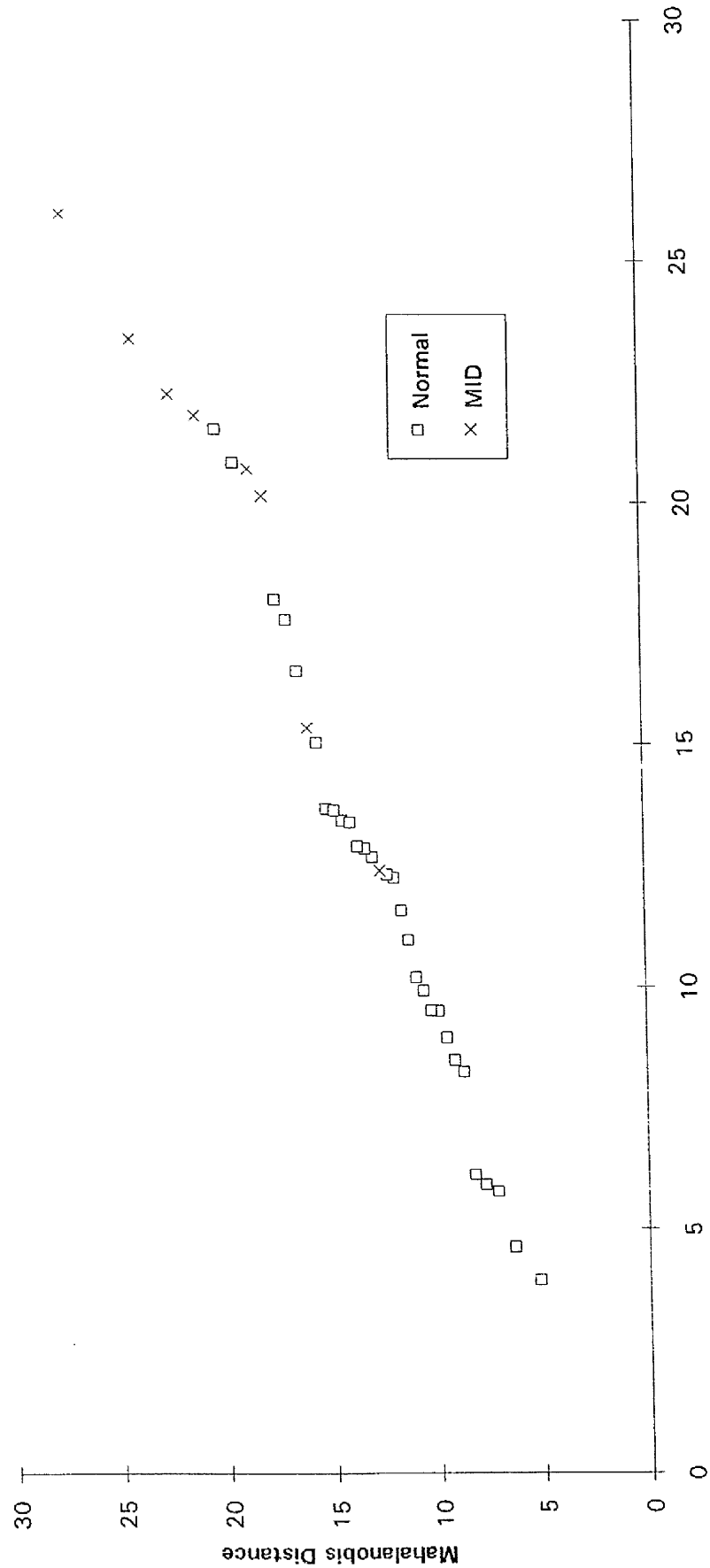


FIGURE 8.3 Probability-plot for screening example data

In terms of the data this is a problem of statistical outlier detection. One approach to outlier detection is through the use of probability-plots. Suppose we took the group of MID subject's (N=8) data and, for the moment, assumed their diagnosis was unknown. A probability-plot of mahalanobis distances, including normals (N=29) and these cases together, is given in Figure 8.3. This suggests that the majority of the unknown cases are similar to the normal group, except possibly for the two cases at the upper extreme of the plot.

8.2 PROBLEMS OF MEASUREMENT

In any measurement process, the data may be subject to several sources of variation, such that repeated observations of the same variable(s) for a subject under similar conditions, may not necessarily yield identical data. This can be seen from ROI data observed in the Alzheimer study. In a small study of reproducibility ten DAT cases had images taken on two occasions one week apart. Figure 8.4 shows the ROI data for one of the cases showing the profiles to be reasonably similar. For various reasons this may not always be the case.

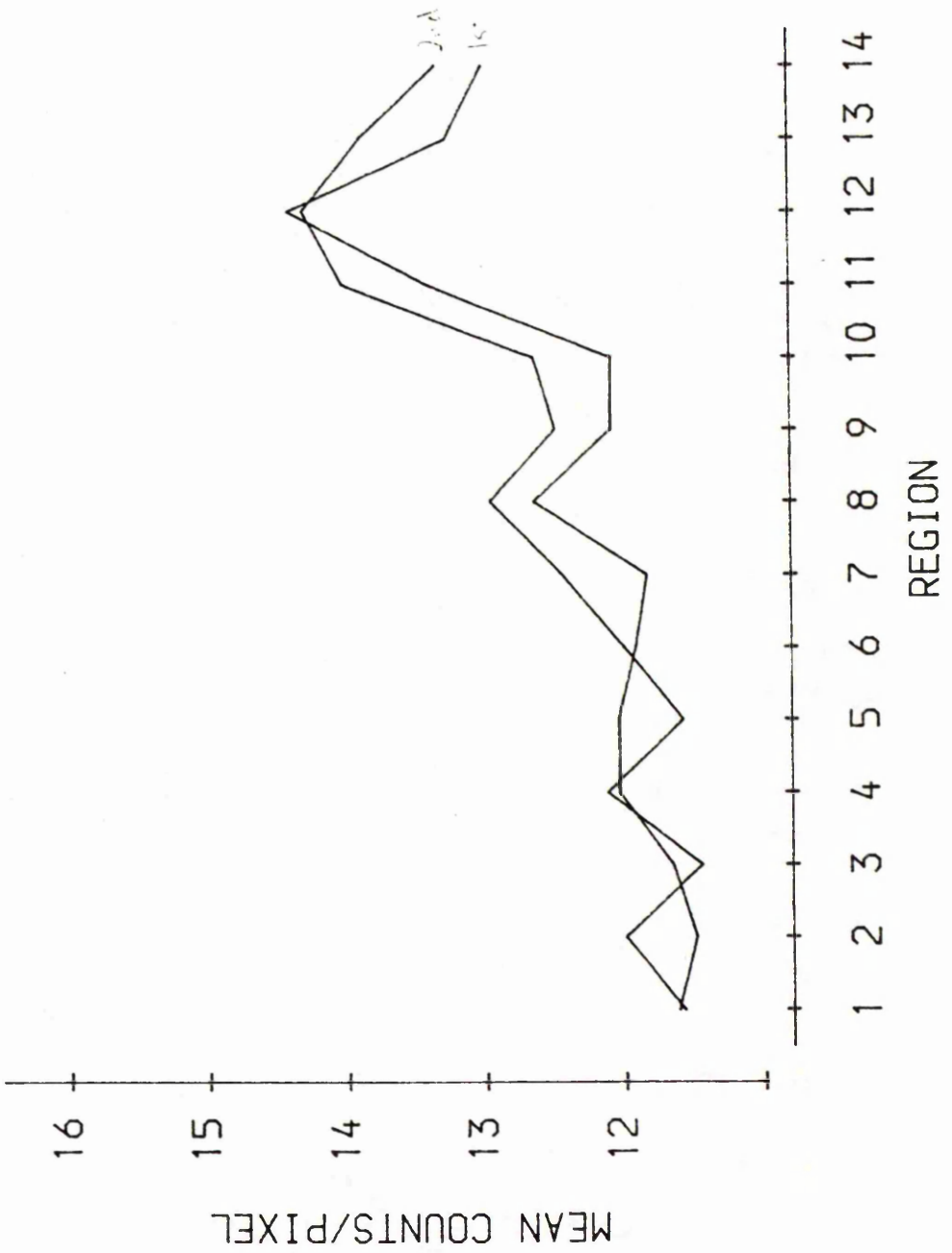
8.2.1 Error propagation in ROI data

As may be apparent from the description of the measurement process in chapter 1, there several stages

involved in ROI data acquisition. These are illustrated in Figure 8.5. The radition data is observed by each detector of the imaging modality at stage A. A tomographic algorithm is then applied to the count data to construct a 2-dimensional or 3-dimensional dataset of pixel or voxel data at stage B. This is then converted into a map of radiation intensity by thresholding the data into grey levels (or an equivalent colour range) which results in the sort of images seen throughout this thesis. In PET imaging, the pixel dataset may undergo an additional transformation wherein pharmacokinetic models are used to translate the pixel data into estimates of the biological parameter of interest in meaningful biological units. The final stage is ROI data extraction. As we have seen this involves identification of regions of interest from the image (Stage C). In the Alzheimer project this is done manually, requiring the expertise of a trained technician to select the appropriate slice/image to use (i.e. `standard' and `upper') and then locate the appropriate anatomical ROI. Recent advances in PET imaging have made it possible to superimpose a structural MRI image onto the functional image (Evans, 1991). This should help reduce any subjectivity and operator bias inherent in the completely manual approach.

A number of factors can be identified which contribute to the measurement error of the ROI data. Some of the factors affecting stage A include diffferences due to head position, head movement, the axial tilt of the image, the physics of radiation detection and the length

figure 8.4 Plot of individuals ROI data taken from scans one week apart.



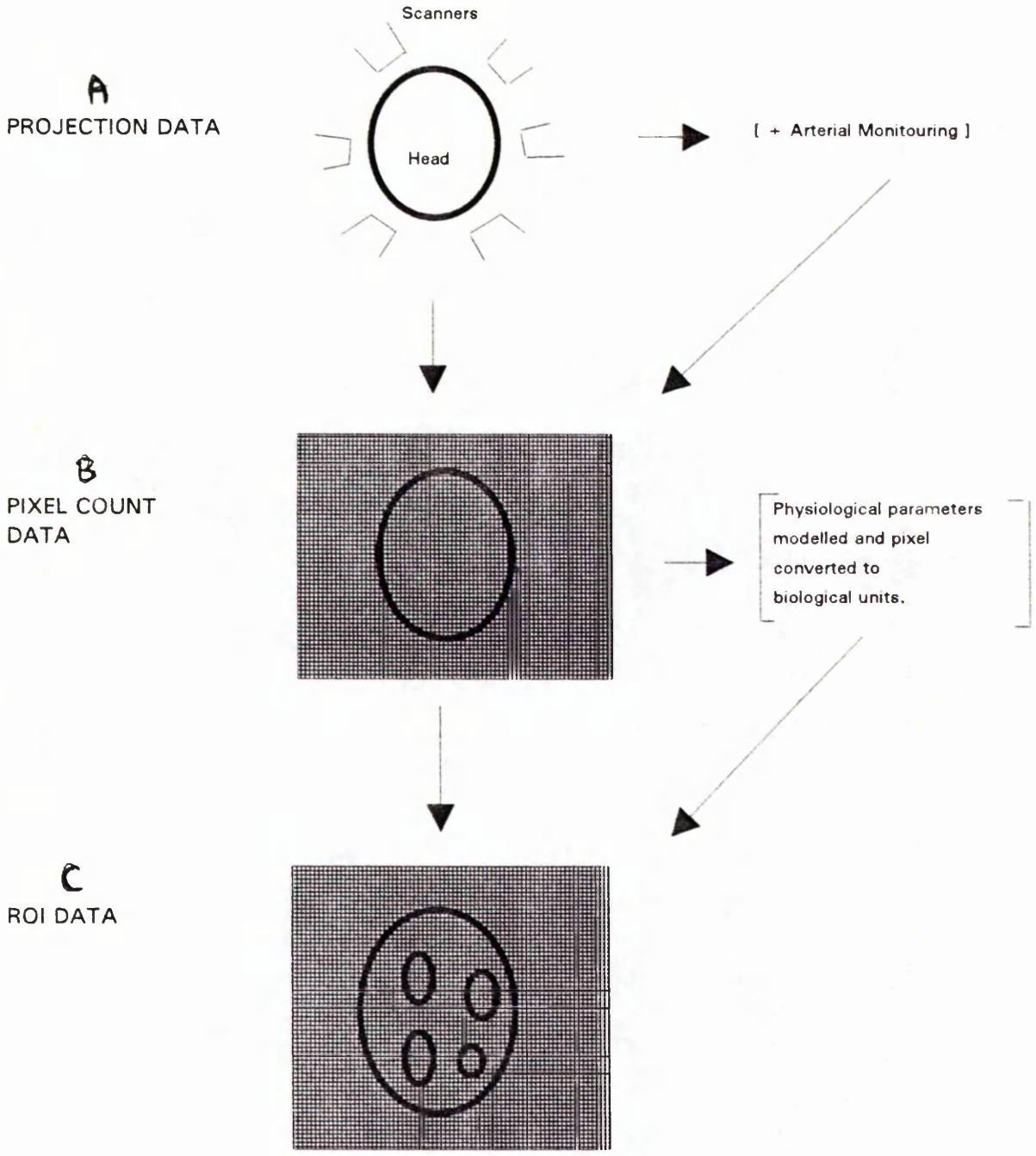

A
PROJECTION DATA
Scanners
Head
[+ Arterial Monitouring]
B
PIXEL COUNT
DATA
Physiological parameters
modelled and pixel
converted to
biological units.
C
ROI DATA

FIGURE 8.5

of time of imaging session. Stage B factors include the type of reconstruction algorithm used, the attenuation (absorption of radiation in tissue) and the scatter corrections used. Finally, at stage C the identification of the anatomical location of the ROIs, the ROI size itself and the ROI summary measure are some factors affecting the reproducibility of the data. For an extensive overview see Mazziotta and Koslow (1987).

A large part of the anatomical location/alignment difficulties are due to the positioning of the head within the scanners. This obviously has an impact on the technicians ability match slices between individuals and in the same individual on different occasions. This may account for the differences between subjects found in section 7.2 - where the size of the posterior ventricles, relative to anterior ventricles was noticeably different between cases A & B and cases C & D.

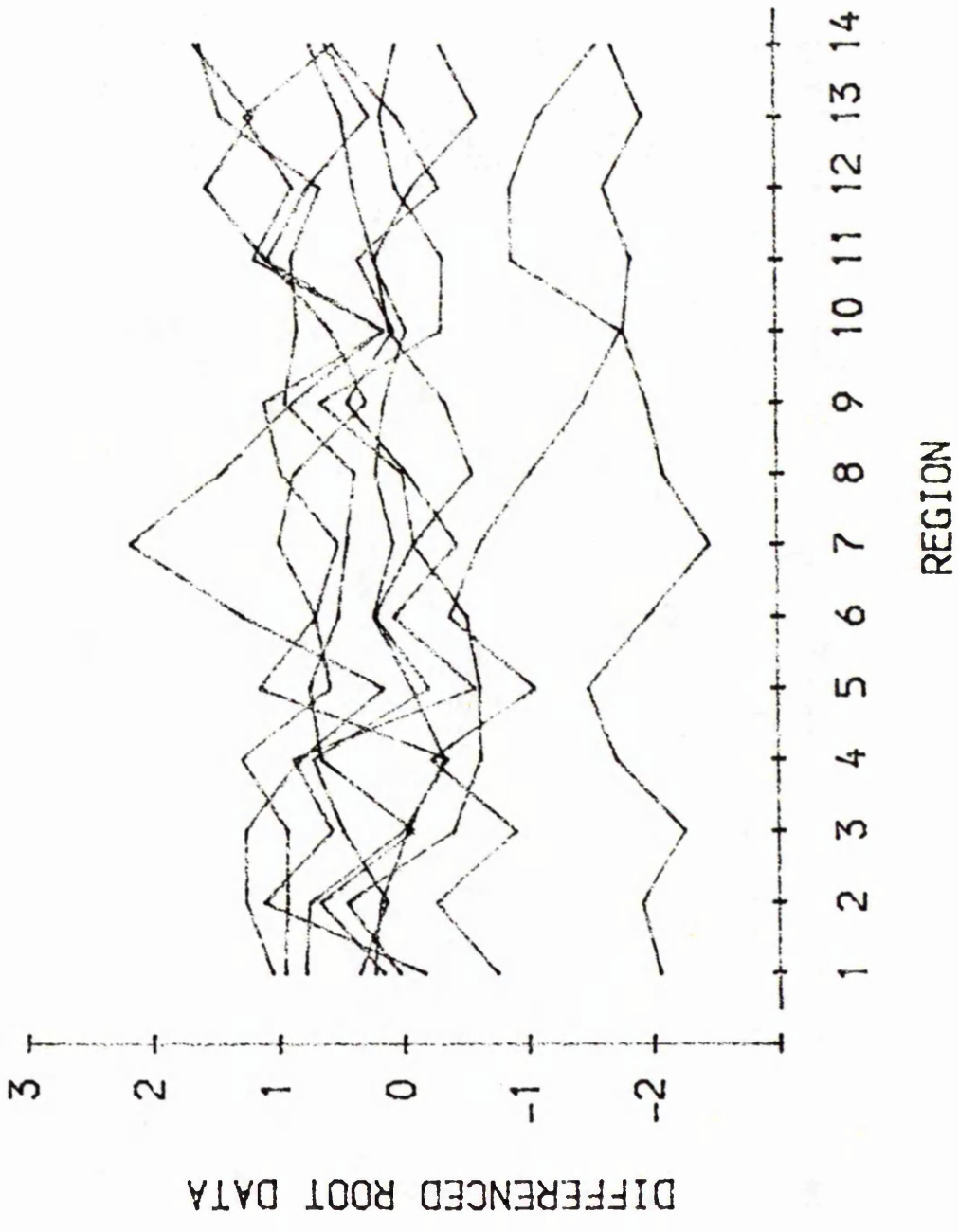
8.2.2 Reprocubility of ROI data

Investigations into the effects of some of these factors have included the study of variation between constructed images (Hoffman et al, 1991; Thompson, 1991) and between reconstruction algorithmns (Strother et al, 1991). A number of small studies have been carried out into the reproducibility of the SPECT ROI data used in the Alzheimer project, involving simple scan/rescan paradigms with and without movement between scan. One of these studies was referred to in section 8.2 and will be

the subject of study in the rest of this section.

As a first step in the analysis of these data, the pairs of ROI data were differenced (Occasion 2-Occasion 1). Figure 8.6 gives a plot of the ten differenced profiles. For some of the cases the patterns do not appear to have been particularly well reproduced; these cases exhibiting large peaks or troughs in some regions. In order to check whether there had been any systematic changes between scanning sessions a univariate repeated measures analysis was carried out on these data comparing regional (differenced) means. Region and hemisphere factors were included in this analysis. The results are shown in Table 8.3. None of the F-tests are significant at the 5% level and hence there is insufficent evidence of any systematic change in profiles between occasions. For each of the tests it was noted that the sphericity assumption (see section 5.5.2) was reasonable; for hemispheric sums p=0.1471 and for hemispheric difference p=0.2135. Thus, sample variance estimates of pairwise differences (of hemispheric sums and hemispheric differences separately) are similar. However, it is interesting to note that the sphericity assumption did not hold for the 14-dimensional data (p=0.0000). We investigated this visually by constructing a two dimensional summary plot using the MDS technique described in section 5.4.3. Pairwise sample variances of differences were calculated between all 14 regions and non-metric multidimensional scaling used to fit a two dimensional solution to the data. As previously, the

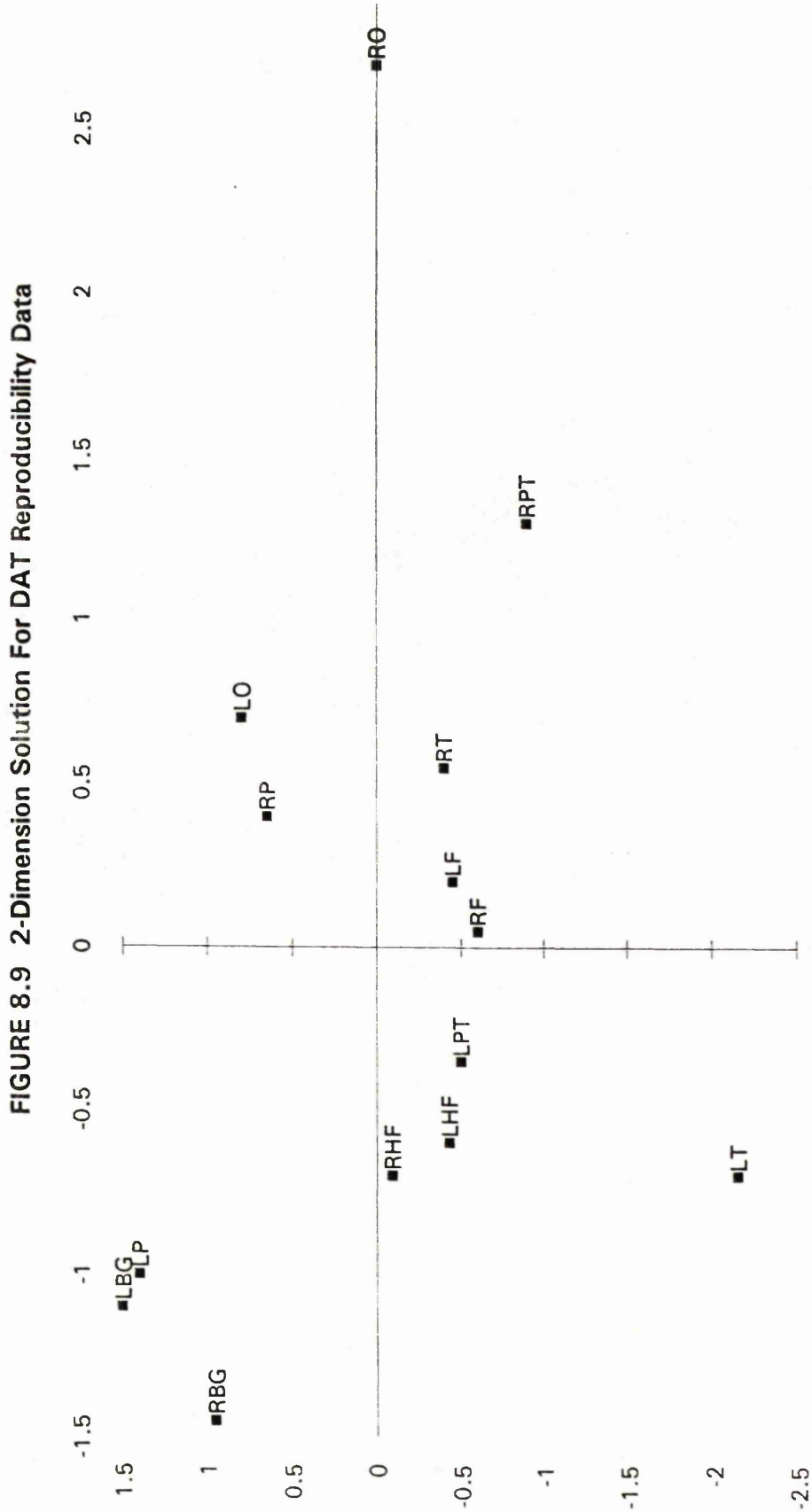
Figure 8.6 Plot of (occasion 1-occasion 2) ROI for DAT subjects.



SOURCE	D.F.	M.S.	F	Unad-P	GG	HF
Between Subject						
Mean	1	0.774	0.07	0.8027		
Error	9	11.694				
Within Subject						
Region	6	0.334	1.52	0.1877	0.2269	0.1969
Error	54	0.219				
Hemisphere	1	0.050	0.27	0.6181		
Error	9	0.187				
Region x Hemisphere	6	0.079	0.85	0.5349	0.4767	0.5133
Error	54	0.093				

TABLE 8.3 : Repeated measures ANOVA analysis of reproducibility data. Adjusted p-values are calculated using Greenhouse-Geisser (GG) and Huyn-Feldt (HF) methods.

Figure 8.7 2-Dimensional solution for (occasion 1-occasion 2) reproducibility data.



ALSCAL program in SPSS-X was used. The two dimensions are plotted in figure 8.7. We can see that the smaller distances are between right and left regions, while larger differences can be seen between different regions. This suggests that bilateral ROI data exhibits less variability between scans than variation between different regions in the same or in different hemispheres.

8.2.3 Further considerations on the use of ROI measurements

Throughout this thesis we concentrated entirely on the analysis of regional SPECT data in terms of the average counts/pixel. It is possible that in certain problems alternative measures of regional activity may more appropriately reflect the biological process under study. As mentioned in section 3.3, the ROI approach to data extraction is effectively the level of resolution at which we can reasonably compare image data across different subjects. However, within regions sub-strates of activity will be present. In order to measure or characterise this, the sample variance between pixel values within the region may be a useful measure. This could be considered as providing a measure of 'texture'. We might even consider using the within-subject rank regional activity for analysis. At the very least this should be robust to some of the factors discussed in this chapter.

8.3 BRAIN ACTIVATION STUDIES

A special type of experimental study, which is becoming increasingly popular to study many of the problems described in this thesis, is the brain activation study. In such studies, the brain is activated or stimulated during the imaging session, so that the active functional state can be compared with the unactivated functional state. The contrast between pre- and post- stimulated images can be be used to identify local and/or global changes in functional profiles which can be considered as anatomical and functional correlates of the form of activation. Activation for these studies is mediated in a number of different ways. Simple cognitive or sensorimotor tasks are common examples (Friston et al 1991; Barker, 1990). A Another form of stimulation includes the intervention of drugs during or prior to imaging, for studying drug/functional metabolism interactions (Heiss, 1988).

The motivation behind such an approach can readily be appreciated using as an example the study of correlation between different brain datasets in chapter seven. In the example comparing imaging data to neuropsychological data (section 7.1), it could be strongly argued that we weren't comparing like with like. The neuropsychological ability of the individuals was being compared to functional activity data in the resting condition. Kennedy (1988) has argued the point vigorously and has suggested that activation studies are the only

worthwhile approach for correlating cognitive function with higher cortical activity.

In general, activation studies involve pre- and post-stimulation imaging of an individual, although there are examples of the use of activation studies for between group comparisons (Bartlett, 1987; Heiss, 1988). The typical design of an activation study is shown in Figure 8.2.

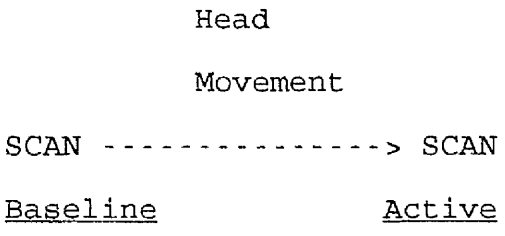


Figure 8.2 Design of a typical activation study.

Because of experimental design issues, e.g. order effects and reallignment difficulties between scans, variations on the basic design have been considered. Friston (1991 and Barker (1988) randomnly assigned the order of rest/stimulation scans. Hunter et al. (1990) used a split-dose approach thus allowing the study of activation without movement between scans.

The data from such studies will consist of an unactivated vector of activity and an activated vector of activity. In the case of ROI data, a simple approach to

analyse the data would be to use the approach to the analysis of paired data used in section 8.2.2. The test of equality of regional means on the differences would indicate whether any change in the ROI pattern had resulted from the form of stimulation. Another approach that may be worth considering here, for the ROI data, is the use of the 'textural' measure of section 8.2.3. This may be more sensitive to subtle substrate changes than the mean counts.

In most activation studies it is in fact the pixel data which is analysed. Friston and colleagues have developed a method of analysing the pixel images known as statistical parametric mapping. This involves subtracting inmages at the pixel level (by analogy with the ROI differences above) and carrying out an ANCOVA analysis, regressing each pixelated difference on the image mean, and using the test on the intercept to assess the significance of any change to the pixel activity resulting from activation. Visually, a pixel map of these p-values is constructed and converted to an image of statistical significicance (SPM). Obviously, with the number of pixels in an image, this approach involves a large number of tests of significance. Friston et al (1990) applied a simple Bonferroni correction to the pixel images to account take account of this fact. However, the smoothing involved in this approach means that the correlation between pixels and hence test may make this approach too conservative. An alternatve testing scheme is described in Worsley et. al. (1992).

CHAPTER 9

CONCLUSIONS AND FURTHER TOPICS

We have considered many of the statistical problems that arise in the analysis of imaging data, with illustrations using SPECT and PET datasets. The main categories of problems we have considered include the analysis of means on cross-sectional data (chapters 3 and 4) & repeated data (chapter 6) and the study of correlation structures among regional levels of activity (chapter 5) & between different study datasets (chapter 7). We also looked at general discrimination type problems and sources of measurement error with imaging data (chapter 8).

The predominant issue in all the analyses has been the presence of large between and within subject variation in data vectors. Because of the need to normalise data prior to analysis, univariate techniques lead to difficulties when interpreting results. With ANOVA techniques we were able to make global comparisons among group mean profiles and use test for various interactions among regions. However, in many practical situations, the multivariate approach is unlikely to be of much value given that typical experiments often involve as many regions as subjects. The beauty of the univariate ANOVA approach is that we can still use it in circumstances; even if $p>>n$ and the sphericity assumption is violated. This approach was also useful for analysing the repeated scan data. With

sensible transformations of the time vectors we saw that useful hypotheses could easily be tested.

Given the potential use of these data for studying functional couplig it is little surprise that correlation type analyses has figured highly in scientific papers in brain research journals. However, the work of Ford (1986) has shown that simple analysis e.g. pairwise correlation analysis of normalised data, is fraught with many pitfalls; the biological interpretation relating to functional anatomy, being complicated by the data transformation itself - which in some cases imposes artificial constraints on data vectors. This would seem to extend to analyses of correlations between imaging data and other medical datasets. Even with this caveat, statistical correlation analyses can be useful in studying overall patterns and in comparing groups. Formal models decribing the covariance structure should provide a useful framework for simulation studies. The work of Worsley et al (1992) has suggested that a spatial model of the covariance structure would be appropriate for imaging data. Correlating imaging data with datasets from different disciplines is another area where the imposition of a transformation to the ROI vectors may cause difficulties in interpretation.

Needless to say imaging data would be a great asset to the clinician or experimental worker in discrimination type problems. The examples in chapter 8 illustrate that it might well be posible to use these

data for these problems. However, this clearly requires further work and would only be viable if used in conjunction with other data.

There are a number of methodological issues which surround the use of imaging techniques. For one, not enough is probably known about the data issues in the measurement process itself, indicated by some of the problems described in chapter eight. Mazziotta & Koslow (1986) discuss these issues at great length.

While at an early stage of development, brain activation techniques offer substantial opportunities for studying some of the problems described in this thesis, without some of the associated problems intimated above.

REFERENCES

AITCHISON, J. (1986). The Statistical Analysis of Compositional Data. Chapman and Hall, London.

ANDERSON, T.W. (1958). An introduction to multivariate statistical analysis. John Wiley, New York.

ATKINSON, K.E. (1978). An introduction to numerical analysis. John Wiley, New York.

BARCLAY, L., ZEMCOV, A., BLASS, J.P., McDOWELL, F. (1984). `Rates of decrease of cerebral blood flow in progressive dementias'. Neurology, 34, 1555-1560.

BARTLETT, E.J., BROWN, J.W., WOLF, A.P., BRODIE, J.D., (1987). `Correlations between glucose metabolic rates in brain regions in healthy male adults at rest and during language stimulation'. Brain and Language, 32, 1-18.

BARTLETT, E.J., BRODIE, J.D., WOLF, A.P., CHRISTMAN, D.R., LASKA, E., MEISSNER, M. (1988). `Reproducibility of cerebral glucose metabolic measurements in resting human subjects'. Journal of Cerebral Blood Flow and Metabolism, 8, 502-512.

BMDP (1985). W.J. DIXON (Ed), Statistical Software, University of California Press, California.

BOX, G.E.P. (1949). `A general distribution theory for a class of likelihood criteria'. Biometrika, 36, 317-346.

BURNS, A., PHILPOT, M.P., COSTA, D.C., ELL, P.J., LEVY, R. (1989). `The investigation of Alzheimer's disease with single photon emission tomography'. Journal of

Neurology, Neurosurgery and Psychiatry, 52, 248-253.

CLARK, C., KESSLER, R., BUCHSBAUM, M., MARGOLIN, R., HOLCOMB, H., (1984). `Correlational methods for determining coupling of regional glucose metabolism: a pilot study. Biological Psychiatry, 19, 663-678.

CLARK, C., CARSON, R., KESSLER, R., MARGOLIN, R., BUCHSBAUM, M., DELISI, L., KING, C., COHEN, R. (1985). `Alternative statistical models for the examination of clinical positron emission tomography/fluorodeoxyglucose data'. Journal of Cerebral Blood Flow and Metabolism, 5, 142-150.

CLARK, C.M., AMMANN, W., MARTIN, W.R.W, TY, P., HAYDEN, M.R. (1991). `The FDG/PET methodology for early detection of disease onset: A statistical model'. Journal of Cerebral Blood Flow and Metabolism, 11, A96-A102.

CRITCHLEY, F., FORD, I. (1982). `Interval estimation in discrimination: The multivariate normal equal covariance case'. Biometrika, 72, 109-116.

CRITCHLEY, F., FORD, I., RIJAL, O. (1988). `Interval estimation based on the profile likelihood: Strong Lagrangian theory with applications to discrimination'. Biometrika, 75, 21-28.

CROWDER, M.J., HAND, D.J. (1990). Analysis of Repeated Measurements. Chapman and Hall, London.

DAVIDSON, M.L. (1972). `Univariate versus multivariate tests in repeated-measures experiments'. Psychological Bulletin, 77, 446-452.

DIGGLE, P.J. (1983). Statistical analysis of spatial point patterns. London Academic Press.

DEWAR, D., WYPER, D., GRAHAM, D.I., MCCULLOCH, J. (1991). `Regional cerebral blood flow and quantitative neuropathology in autopsy-verified cases of Alzheimer's disease'. Journal of Cerebral Blood Flow and Metabolism, 11, Supplement 2, S18.

DUARA, R., BARKER, W.W., PASCAL, S., BRUCE-GREGORIOS, J., NOREMBERG, M., BOOTHE, T. (1991). `Lack of correlation of regional neuropathology to the regional PET metabolic deficits in Alzheimer's disease (AD)'. Journal of Cerebral Blood Flow and Metabolism, 11, Supplement 2, S19.

EFRON, B., TIBSHIRANI, R., (1986). `Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy'. Statistical Science, Vol. 1, No. 1, 54-77.

ELL, P.J. (1990). `Single photon emission computed tomography (SPET) of the brain'. Journal of Neuroscience Methods, 34, 207-217.

ENGEL, J., KUHL, D.E., PHELPS, M.E., CRANDALL, P.H. (1982). `Comparative localization of epileptic foci in partial epilepsy by PCT and EEG'. Annals of Neurology, 12, 529-537.

EVANS A.C., MARRETT, S., TORRESCORZO, J., KU, S., COLLINS, L. (1991). `MRI-PET correlation in three dimensions using a volume of interest (VOI) atlas'. Journal of Cerebral Blood Flow and Metabolism, 11, A69-A78.

FORD, I. (1986). `Short communication, Confounded Correlations: Statistical limitations in the

analysis of interregional relationships of cerebral metabolic activity'. Journal of Cerebral Blood Flow and Metabolism, 6, 385-388.

FOX, P.T., MINTUN, M.A., REIMAN, E.M., RAICHLE, M.E. (1988). `Enhanced detection of focal brain responses using intersubject averaging and change distribution analysis of subtracted PET images'. Journal of Cerebral Blood Flow and Metabolism, 8, 642-653.

FRISTON, K.J., PASSINGHAM, R.E., NUTT, J.G., HEATHER, J.D., SAWLE, G.V., FRACKOWIAK, R.S.J. (1989). `Localisation in PET images: Direct fitting of the intercommissural (AC-PC) line'. Journal of Cerebral Blood Flow and Metabolism, 9, 690-695.

FRISTON, K.J., FRITH, C.D, LIDDLE, P.F., LAMMERTSMA, A.A., DOLAN, R.D., FRACKOWIAK, R.S.J. (1990). `The relationship between local and global changes in PET scans'. Journal of Cerebral Blood Flow and Metabolism, 10, 458-466.

FRISTON, K.J., FRITH, C.D, LIDDLE, P.F., FRACKOWIAK, R.S.J. (1991). `Comparing functional (PET) images: The assessment of significant change'. Journal of Cerebral Blood Flow and Metabolism, 11, 690-699.

GEMMELL, H.G., SHARP, P.F., BESSON, J.A.O., CRAWFORD, J.R., EBMEIER, K.P., DAVIDSON, J., SMITH, F.W., (1987). `Differential diagnosis in dementia using cerebral blood flow agent 99mTc HM-PAO: A SPECT study'. Journal of Computer Assisted Tomography, 11(3), 398-402.

GNANADESIKAN, R. (1977). Methods for statistical data

analysis of multivariate observations. John Wiley, New York.

GREENHOUSE, S.W. and GEISSER, S. (1959). `On methods in the analysis of profile data'. Psychometrika, 24, No. 2, June.

HAND, D.J. (1981). Discrimination and Classification. John Wiley & Sons.

HAXBY, J.V., DUARA, R., GRADY, C.L., CUTLER, N.R., RAPOPORT, S.I. (1985). `Relations between neuropsychological and cerebral metabolic asymmetries in early Alzeimer's disease'. Journal of Cerebral Blood Flow and Metabolism, 5, 193-200.

HOFFMAN, E.J., DUFFY CUTLER, P., GUERRORO, T.M., DIGBY, W.M., MAZZIOTTA, J.C., (1991). `Assessment of accuracy of PET utilizing a 3-D phantom to simulate the activity distribution of [18F]fluorodeoxyglucose uptake in the human brain'. Journal of Cerebral Blood Flow and Metabolism, 11(2), A17-A25.

HORWITZ, B., DUARA, R., RAPOPORT, S.I. (1984). `Intercorrelations of glucose metabolic rates between brain regions: Application to healthy males in a state of reduced sensory input'. Journal of Cerebral Blood Flow and Metabolism, 4, 484-499.

HORWITZ, B., DUARA, R., RAPOPORT, S.I. (1985). `Glucose use correlations: a reply'. Journal of Cerebral Blood Flow and Metabolism, 5, 620-621.

HORWITZ, B., GRADY, C.L., SCHLAGETER, N.L., DUARA, R., RAPOPORT, S.I. (1987). `Intercorrelations of regional cerebral metabolic rates in Alzheimer's disease'.

Brain Research, 407, 294-306.

HORWITZ, B. (1990). `Functional interactions in the brain: use of correlations between regional metabolic rates'. Journal of Cerebral Blood Flow and Metabolism, 11, A114-A120.

HUYN, H. and FELDT, L.S. (1976). `Estimation of the Box correction for degrees of freedom from sample data in randomised block and split-plot designs'. Journal of Educational Statistics, 1, 69-82.

JACOB HUFF, F., AUERBACH, J., CHAKRAVARTI, A., BOLLER, F. (1988). `Risk of dementia in relatives of patients with alzheimer's disease', Neurology, 38, 786-790.

JAGUST, W.J., BUDINGER, T.F., REED, B.R., (1987). `The diagnosis of dementia with single photon emission computed tomography'. Arch Neurol, 44, 258-262.

JAGUST, W.J., FRIEDLAND, R.P., BUDINGER, T.F., KOSS, E., OBER, B. (1988). `Longitudinal studies of regional cerebral metabolism in Alzheimer's disease'. Neurology, 38, 902-912.

KATZMAN, R. (1976). `The prevalence and malignancy of Alzheimer disease: a major killer'. Archives of Neurology, 33, 217-218.

KELLY, P.A.T., FORD, I. McCULLOCH, J., (1982). `Effect of Apomorphine on the relationship between local cerebral glucose utilisation and local cerebral blood flow (with an appendix on its statistical analysis)'. Journal of Cerebral Blood Flow and Metabolism, 2, 487-499.

KENWARD, M.G. (1987). `A method for comparing profiles of

repeated measurements'. The Journal of the Royal Statistical Society, Series C, 36, No. 3, 296-308.

KHACHATURIAN, Z. (1985). `Diagnosis of Alzheimer's disease'.Archives of Neurology, 42, 1097-1105.

KRUSKAL (1964). `Nonmetric multidimensional scaling: a numerical method'. Psychometrika, 29, 115-129.

KRZANOWSKI, W.J. (1988). Principles of Multivariate Analysis:a user's perspective. Oxford Science Publications.

KUSCHINSKY, W., SUDA., S., SOKOLOFF, L. (1981). `Local cerebral glucose utilisation and blood flow during metabolic acidosis'. American Journal of Physiology, 241, H772-H777.

McCORMACK, A.G. (1991). Some Problems in the Statistical Analysis of Neuroimages. Phd Thesis, University of Glasgow.

MANLY, B.F.J., RAYNER, J.C.W. (1987). `The comparison of sample covariance matrices using likelihood ratio tests'. Biometrika, 74, 4, 841-847.

MARDIA, K.V., KENT, J.T., BIBBY, J.M. (1979). Multivariate Analysis. Academic Press, London.

MARTIN, G.M., SCHELLENBERG, G.D., WIJSMAN, E.M., BIRD, T.D. (1990). `Dominant susceptibility genes'. Nature, Vol 347, 124.

MAZZIOTTA, J.C., KOSLOW, S.H., (1987). `Assessment of the goals and obstacles in data acquisition and analysis from emission tomography: Report of a series of international workshops'. Journal of Cerebral Blood Flow and Metabolism, 7, S1-S31.

METTER, E.J., RIEGE, W.H., KUHL, D.E., PHELPS, M.E. (1984a). `Cerebral metabolic relationships for selected brain regions in healthy adults'. Journal of Cerebral Blood Flow and Metabolism, 4, 1-7.

MOELLER, J.R., STROTHER, S.C., SIDTIS, J.J., ROTTENBERG, D.A. (1987). `Scaled subprofile model: A statistical approach to the analysis of functional patterns in positron emission tomographic data'. Journal of Cerebral Blood Flow and Metabolism, 7, 649-658.

MOLSA P.K., MARTILA, F.J., RINNE, U.K. (1982). `Epidemiology of dementia in a Finnish population'. Acta Neurologica Scandinavica, 65, 541-552.

MONTALDI, D., BROOKS, D.G., McCOLL, J.H., WYPER, D., PATTERSON, J., BARRON, E., McCULLOCH, J. (1990). `Measurements of regional cerebral blood flow and cognitive performance in Alzheimer's disease'. Journal of Neurology, Neurosurgery and Psychiatry, 53, 33-38.

MORGAN, W.A. (1939). `A test for the significance of the difference between two variances in a sample from a normal bivariate population'. Biometrics, 31, 13-19.

PERANI, D., DI PIERO, V., LUCIGNANI, G., GILARDI, M.C., PANTANO, P., ROSETTI, C., POZZILLI, C., GERUNDINI, P., FAZIO, F., LENZI, G.L. (1988). `Remote effects of subcortical cerebrovascular lesions: A SPECT cerebral perfusion study'. Journal of Cerebral Blood Flow and Metabolism, 8, 560-567.

PERL, D. (1989). Article in December 18 edition of Newsweek.

PERLMUTTER, J.S., HERSCOVITCH, P., POWERS, W.J., FOX, P.T., RAICHLE, M.E. (1985). `Standardised mean regional method for calculating global positron emission tomographic measurements'. Journal of Cerebral Blood Flow and Metabolism, 5, 476-480.

PERLMUTTER, J.S., POWERS, W.J., HERSCOVITCH, P., FOX, P.T., RAICHLE, M.E. (1987). `Regional asymmetries of cerebral blood flow, blood volume, and oxygen utilisation and extraction in normal subjects'. Journal of Cerebral Blood Flow and Metabolism, 7, 64-67.

PHELPS, M.E., MAZZIOTTA, J.C., SCHELBERT, H.R. (1979). Positron Emission Tomography: Principles and applications for the brain and heart. Raven Press, New York.

PITMAN, E.J.G. (1939). `A note on normal correlation'. Biometrics, 31, 9-12.

PROHOVNIK, ., MAYEUX, R., SACKEIM, H.A., SMITH, G., STERN, Y., ALDERSON, P.O. (1988). `Cerebral perfusion as a diagnostic marker of early Alzheimer's disease'. Neurology, 38, 931-937.

RHODES, C.G., WISE, R.J.S., GIBBS, J.M, FRACKOWIAK, R.S.J., HATAZAWA, J., PALMER, A.J., THOMAS, D.G.T, JONES, T. (1983). `In vivo disturbance of the oxydative metabolism of glucose in human cerebral gliomas'. Annals of Neurology, 14, 614-626.

RIGBY, R.A. (1982). `A credibility interval for the probability that a new observation belongs to one of two multivariate normal populations'. Journal of the

Royal Statistical Society, Series B, 44, 2, 212-220.

RISBERG, J. (1985). `Cerebral blood flow in dementias'. Danish Medical Bulletin, 32(suppl 1).

RODRIGUEZ, G., COGORNO, P., GRIS, A., MARENCO, S., MESITI, C., NOBILI, F., ROSADINI, G. (1989). `Regional cerebral blood flow and anxiety: A correlation study in neurologically normal patients'. Journal of Cerebral Blood Flow and Metabolism, 9, 410-416.

ROTH, M., TYM, E., MOUNTJOY, C.Q., HUPPERT, F.A., HENDRIE, H., VERMA, S., GODDARD, R. (1986). `CAMDEX A standardised instrument for the diagnosis of mental disorder in the elderly with special reference to the early detection of dementia'. British Journal of Psychiatry, 149, 698-709.

SCHAAFSMA, W., VAN VARK, G.N. (1979). `Classification and discrimination problems with applications'. IIa., Statist. Neer., 33, 91-126.

SCHOENBERG, B.S. (1986). `Epidemiology of Alzheimer's disease and other dementing illnesses'. Journal of Chronic Diseases, Vol 39, 12, 1095-1104.

SHEDLACK, K.J., HUNTER, R., WYPER, D., MCLUSKIE, R., FINK, G., GOODWIN, G.M, (1989). `The pattern of cerebral activity underlying verbal fluency shown by split-dose single photon emission tomography (SPET) in normal volunteers'. Soc Neurosci, 15, 27- ??? .

SILVERMAN, B.W., JONES, M.C., WILSON, J.D. and NYCHKA, D.W. (1990). `A smoothed EM approach to indirect estimation problems, with particular reference to stereology and emission tomography'. Journal of the

Royal Statistical Society, Series B, 52, 271-324.

SMITH, F.W., BESSON, J.A.O., GEMMELL, H.G., SHARP, P.F., (1988). `The use of Tecnetium-99m-HM-PAO in the assessment of patients with dementia and other neuropsychiatric conditions'. Journal of Cerebral Blood Flow and Metabolism, 8, S116-S122.

SNEDECOR, G.W., COCHRAN, W.G. (1980). Statistical Methods. 7th Edition, Iowa State University Press, Ames, Iowa.

SONCRANT, T.T., HORWITZ, B., HOLLOWAY, H.W., RAPOPORT, S.I. (1986). `The pattern of functional coupling of brain regions in the awake rat'. Brain Research, 369, 1-11.

SPSSX. Statistical software, SPSS Inc., USA.

STROTHER, S.C., LIOW, J-S., MOELLER, J.R., SIDTIS, J.I., DHAWAN, V., ROTTENBERG, D.A., (1991). `Absolute quantitation in neurological PET: Do we need it'. Journal of Cerebral Blood Flow and Metabolism, 11(2), A3-A16.

TESTA, H.J., SNOWDEN, J.S., NEARY, D., SHILEDS, R.A., BURJAN, A.W.I., PRESCOTT, M.C., NORTHEN, B., GOULDING, P., (1988). `The use of [99mTc]-HM-PAO in the diagnosis primary degenrative dementia'. Journal of Cerebral Blood Flow and Metabolism, 8, S123-S126.

THOMPSON, C.J., (1991). `The effects of collimation on PET image noise due to scatter, random counts, and deadtime'. Journal of Cerebral Blood Flow and Metabolism, 11(2), A31-A37.

TYLER, J.L., STROTHER, S.C., ZATORRE, R.J., ALIVISATOS,

B., WORSLEY, K.J., DIKSIC, M., YAMAMOTO, Y.L. (1988). `Stability of regional cerebral glucose metabolism in the normal brain measured with positron emission tomography'. Journal of Nuclear Medicine, 29, 631-642.

VARDI, Y., SHEPP, L.A., KAUFMAN, L. (1985). `A statistical model for positron emission tomography (with comments)'. Journal of the American Statistical Association, 80, 8-37.

VOLKOW, N.D., BRODIE, J.D., WOLF, A.P., GOMEZ-MONT, F., CANCRO, R., VAN GELDER, P., RUSSEL, J.A.G., OVERALL, J. (1986). `Brain organisation in schizophrenia'. Journal of Cerebral Blood Flow and Metabolism, 6, 441-446.

WIESZNIEWSKY (1988).

WORSLEY, K.J., EVANS, A.C, STROTHER, S.C., TYLER, J.L. (1991). `A linear spatial correlation model with applications to positron emission tomography'. Journal of the American Statistical Association, 86, 55-67.

WORSLEY, K.J., EVANS, A.C, MARRETT, S., NEELIN, P., (1992). `Determining the number of statistically significant areas of activation in subtracted activation studies from PET'. In submission.