STATISTICAL ISSUES
IN
*f*UNCTIONAL BRAIN MAPPING

Andrew Peter Holmes

Submitted to the
University of Glasgow
for the Degree of
Doctor of Philosophy

Department of Statistics,
University of Glasgow,
December 1994

To Mum, Dad, and Rachel.

Abstract

Using Positron Emission Tomography (PET) it is possible to obtain quantitative images of regional cerebral blood flow, indicative of regional neuronal activity. This is used to examine the function of the brain through designed experiments. The simplest such *functional mapping experiment* is the *simple activation study*, which aims to locate the brain loci responsible for a certain function, by scanning under two conditions differing only in that function. These studies generate small numbers of observations of extremely high dimension, each data point being a three-dimensional image. The statistical analysis of such data sets requires new statistical techniques for significance testing, and it is this problem which this thesis addresses.

Chapter 1 establishes the background to this work, giving a fairly comprehensive layman's description of PET and functional brain mapping.

In the absence of apriori information regarding the location of a particular function, analysis of activation experiments proceeds at the voxel (pixel) level. For each voxel a model must be assumed for the data, and a null hypothesis of "no activation" expressed in terms of the model parameters. An added complication is the presence of global differences in cerebral blood flow between subjects. Computation of a statistic indicating evidence against the null hypothesis at each voxel, gives a statistic image. Model selection and the formation of statistic images is the subject of chapter 2. Particular attention is given to the two most popular models, namely Friston's ANCOVA and that of the *t*-statistic formed from subject difference images, with global changes removed by proportional scaling. Problems of simultaneous model fitting and the dangers of assuming homoscedascity are considered.

The assessment of the statistic image presents a large multiple comparisons problem. Regions where the statistic image indicates evidence against the null hypothesis must be located, whilst maintaining strong control over familywise Type I error. The current methods for testing statistic images are discussed in detail in chapter 3, focusing on "random field" methods, their assumptions and properties.

In the remaining chapters, three methods developed by the author for assessing simple activation studies are presented. The first of these is a two-stage approach, in which the group of subjects is split into a pilot group and a study group. A small number of regions of interest are identified from the pilot group data, and the study group data assessed over these regions of interest. A simulation study shows this approach to hold some promise. In chapter 5, the testing problem is reformulated as an image segmentation problem, to which empirical Bayesian techniques from statistical image processing are applied. A Markov random field is used to convey prior belief regarding the contiguous nature of activated areas. Here simulation results indicate that the incorporation of prior belief into a single threshold test results in a more conservative (and less powerful) test.

The subject of the last chapter (ch.6), is a non-parametric approach. A multiple comparisons randomisation test is developed for simple activation studies, which is shown to maintain strong control over familywise Type I error. A step-down procedure with strong control is introduced, and computationally feasible algorithms presented. The methods are illustrated on a real PET data set, with a pseudo *t*-statistic formed from subject difference images with a smoothed variance estimate. For the given data set the

continued…

approach is found to outperform many of the parametric methods, particularly with the pseudo *t*-statistic. This, together with the flexibility and guaranteed validity of a non-parametric method, makes the approach very attractive, despite the computational burden imposed. The practicalities of the method are discussed, including extensions to other experimental paradigms, other test statistics, and permutation tests.

This is an applied thesis, aimed at the statistically literate PET researcher. In addition to presenting the author's ideas, it is hoped that this document provides a useful summary and comprehensive critique of existing work, giving enough detail to serve as a useful reference.

Table of Contents

List of Figures and Tables [†]

Chapter 1: PET Background

Chapter 2: Statistic Images

Chapter 3: Current Methods for Testing Statistic Images

[†] Figures and tables have been numbered together.

Chapter 4: Two Stage Testing

Chapter 5: An Empirical Bayesian Approach

Chapter 6: A Non-Parametric Approach

Appendices

Acknowledgements

I would like to thank my supervisor, Professor Ian Ford, for his friendly advice and encouragement throughout the three (and a quarter) years leading to the submission of this thesis, and for providing me with first class resources.

I would also like to record my gratitude to the following people:

> Dr. Karl Friston, of the Wellcome Department of Cognitive Neurology at the Institute of Neurology; for his enthusiasm and interest in this Ph.D., and for a number of very interesting discussions.
>
> Dr. John Watson, who, whilst a research fellow at the Medical Research Council Cyclotron Unit, took an active interest in my work, encouraged me enormously, helped me to understand the methods of analysis used at the time, and provided the "V5" data set analysed at length in this thesis.
>
> Dr. Keith Worsley, of the Statistics Department, McGill University, Montreal; for his speedy responses to emailed questions, and for his advice.
>
> Mr. David Duncan, formerly data manager at the Robertson Centre for Biostatistics in the Statistics Department at the University of Glasgow; for his interest and assistance in all things computing.
>
> Professor Richard Frackowiak, of the Wellcome Department of Cognitive Neurology at the Institute of Neurology, formerly director of the Medical Research Council Cyclotron Unit, Clinical Sciences Section; for making it possible to collaborate with fellows at the unit, and for giving me a job.

On a more practical note, I am indebted to the Engineering and Physical Sciences Research Council (formerly the Science and Engineering Research Council) for their continued funding of my postgraduate studies.

Finally, I would like to thank my parents for their love and support.
Thank you for everything.

Declaration

Unless otherwise stated in the text, the work presented in this thesis is my own.

In attempting to provide a comprehensive review of existing methods, chapters 2 and 3 contain summaries of other peoples work, some slightly reworked. In general the amount of reworking is minimal, and these summaries should be viewed as summaries. However, the evaluation of these methods is my work. Appendix C contains a number of results regarding smoothing of continuous random fields. These are doubtless available elsewhere, but I couldn't find them, and derived them myself (with the exception of result C:6 which is due to Adler, 1981).

The empirical Bayesian approach of chapter 5 was presented orally at Brain PET'93, the first International Symposium on Quantification of Brain Function, held in Akita, Japan. An abstract appears in the *Annals of Nuclear Medicine* (Holmes & Ford, 1993a), and a full paper in the conference proceedings, *Quantification of Brain Function: Tracer Kinetics and Image Analysis in PET* (Holmes & Ford, 1993b).

The randomisation approach of chapter 6 has been accepted for publication in the *Journal of Cerebral Blood Flow and Metabolism* (Holmes *et al.*, 1995), and has been the subject of invited seminars at University of Pittsburgh Medical Centre PET Facility (September 1993); The Department of Nuclear Medicine, Royal Prince Alfred Hospital, Sydney, Australia (June 1993); and the MRC Cyclotron Unit Clinical Sciences Section, Hammersmith Hospital, London (February 1993).

Andrew P. Holmes.
Ph.D., 1994.

Introduction

Mind and Brain

One of the many exciting areas of current medical research is the study of the living human brain. The brain is the most complex structure in the known universe, and its workings, perhaps rather vainly, have always fascinated man. Current findings and methods are frequently reported in the popular press (Vogue, 1992; The Observer, Sunday 2nd December 1990). The link between the mind and the brain has troubled philosophers for centuries. René Descartes described the mind three centuries ago as an "extra corporeal entity that was expressed through the pineal gland". Modern scientific medicine has taken the study of the mind out of the hands of the philosophers. Armed with powerful research tools, medical researchers have gained considerable knowledge of the brain, recently summarised in a special issue of *Scientific American* (1993). One of these research tools is Positron Emission Tomography.

Positron Emission Tomography

Positron Emission Tomography (PET) is a medical imaging modality that enables quantitative images of the density of a radioactive tracer in the living body to be obtained. Using radioactive bio-chemicals, measurement of various physiological, metabolic and biochemical processes within a living subject is possible.

Functional mapping

Of particular interest are PET experiments where regional blood flow is measured repeatedly in the human brain. Regional cerebral blood flow is linked to regional neuronal activity, so these images are indicative of brain activity. This discovery has led to the use of designed experiments to examine the function of the human brain. The simplest such *functional mapping* experiment is the *activation study*, which aims to locate the brain region responsible for a certain function. These proceed by acquiring scans under two conditions, which differ in the mental processing required, only in the function of interest. Typically the conditions are simply a "rest" and an "activation" condition.

These studies generate small numbers of observations of extremely high dimension, each data point being a three-dimensional image. The statistical analysis of such data sets require new statistical techniques for significance testing, and it is this problem which this thesis addresses.

Outline

The thesis covers three areas, background, a review of existing methods, and a presentation of the author's efforts.

Chapter 1 establishes the background to this work, giving a fairly comprehensive overview PET and functional mapping. Covered are: the PET tomograph, the physics of PET and SPECT, annihilation detection, image reconstruction, radio-tracer issues, and functional mapping experiments. The aim of this chapter is to set the scene for the reader unfamiliar with PET.

With the background to the data covered, we turn to the main topic of this research, hypothesis testing for simple activation studies.

For each voxel a model must be assumed for the data, and a null hypothesis of "no activation" expressed in terms of the model parameters. An added complication is the presence of global differences in cerebral blood flow between subjects. Computation of a statistic indicating evidence against the null hypothesis at each voxel, gives a statistic image. Model selection and the formation of statistic images is the subject of chapter 2. Using a real PET activation data set, a simultaneous regression analysis is applied to elucidate suitable ANCOVA models for the single and multiple subject activation experiment. Particular attention is given to the two most popular models, namely Friston's ANCOVA and that of the *t*-statistic formed from subject difference images, with global changes removed by proportional scaling. Problems of simultaneous model fitting and the dangers of assuming homoscedascity are considered. The chapter concludes with example statistic images.

The remainder of the thesis addresses the large multiple comparisons problem of assessing these statistic images.

Regions where the statistic image indicates evidence against the null hypothesis must be located, whilst maintaining strong control over familywise Type I error. In chapter 3, the required theory and terminology is reviewed, and the existing methods expounded in sufficient detail to enable their implementation. The validity of these methods is assessed under idealised conditions with a set of simulated Gaussian statistic images. The popular "random field" methods are considered at length, and their assumptions and properties discussed. The chapter closes with analyses of a PET data set using a variety of methods.

In the final chapters, three approaches developed by the author are presented. The first of these is a two-stage approach, in which the group of subjects is artificially split, post hoc, into a pilot group and a study group. A small number of regions of interest are

identified from the pilot group data. The study group data is then analysed over these regions, rather than over all the voxels individually. The idea is that the substantial reduction in the number of simultaneous tests affords an increase in power over a voxel-by-voxel approach on the whole data set, despite the reduced number of subjects available for thefinal testing. This is examined in a simulation study, where a number of the existing voxel-by-voxel approaches are compared with the two-stage method, applied to sets of two dimensional *t*-statistic images formed from simulated Gaussian subject difference images to which focal signals of various amplitudes have been added.

In chapter 5, the testing problem is reformulated as an image segmentation problem, to which empirical Bayesian techniques from statistical image processing are applied. In particular, a Markov Random Field is used to convey prior belief regarding the contiguous nature of activated voxels. Here, simulation results indicate that the incorporation of prior belief into a single threshold test results in a more conservative (and less powerful) test.

The subject of the last chapter (ch.6), is a non-parametric approach. A multiple comparisons randomisation test is developed for simple activation studies, which is shown to maintain strong control over familywise Type I error. A step-down procedure with strong control is introduced, and computationally feasible algorithms presented. The methods are illustrated on a real PET data set, with a pseudo *t*-statistic image formed using a smoothed variance estimate. For the given data set the approach is found to outperform many of the parametric methods, particularly with the pseudo *t*-statistic. This, together with the flexibility and guaranteed validity of a non-parametric method, makes the approach very attractive, despite the computational burden imposed. The practicalities of the method are discussed, including extensions to other experimental paradigms, other test statistics, and permutation tests.

Each chapter contains its own conclusions section where further work is suggested. Thus, there is no conclusions chapter *per se*.

The references are preceded by appendices containing technical details, results and proofs.

Style

This is an applied thesis, written primarily with the statistically literate PET researcher in mind. For such a reader, it is hoped that in addition to presenting the author's ideas, this document provides a useful summary and a comprehensive critique of existing work, giving sufficient detail to serve as a useful reference.

It is hoped that the resulting level of exposition is not a nuisance to the statistical reader.

Omissions

Inevitably some statistical areas of functional neuroimaging have had to be neglected in this work. The most notable omissions are the *scaled subprofile model* (Moeller *et al.*, 1987, and Moeller & Strother, 1991), and the principal components analyses which are rapidly becoming popular to investigate *functional connectivity* (Friston *et al.*, 1993a&b).

Caveat

Our knowledge of the brain is increasing in leaps and bounds. Whether the mind can ever finally know itself, or whether it will stay a step ahead of its pursuer, like the tortoise pursued by the Achilles in Zeno's paradox,[1] remains a question for the philosophers.

[1]Zeno of Elea, born around 469B.C., was the originator of the "Race Course" paradox. This is the paradox of a runner never reaching his goal, since he must traverse infinitely many successive halves of the distance. Zeno's work is discussed in the Encyclopædia of Philosophy (1967), **7**:369–379 (Collier-Macmillan publishers, London).