STATISTICS IN MEDICINE
Statist. Med. 2004; **23**:315–326 (DOI: 10.1002/sim.1719)

A comparison of classification methods for differentiating fronto-temporal dementia from Alzheimer's disease using FDG-PET imaging

Roger Higdon[1,*,†], Norman L. Foster[2], Robert A. Koeppe[2], Charles S. DeCarli[3], William J. Jagust[3], Christopher M. Clark[4], Nancy R. Barbas[2], Steven E. Arnold[4], R. Scott Turner[2], Judith L. Heidebrink[2] and Satoshi Minoshima[1]

[1] *University of Washington, U.S.A.*
[2] *University of Michigan, U.S.A.*
[3] *University of California Davis, U.S.A.*
[4] *University of Pennsylvania, U.S.A.*

SUMMARY

Flurodeoxyglucose positron emission tomography (FDG-PET) is being explored to determine its ability to differentiate between a diagnosis of Alzheimer's disease (AD) and fronto-temporal dementia (FTD). We have examined statistical discrimination procedures to help achieve this purpose and compared the results to visual ratings of FDG-PET images. The methods are applied to a data set of 48 subjects with autopsy confirmed diagnoses of AD or FTD (these subjects come from a multi-centre collaborative study funded by the National Alzheimer's Coordinating Center). FDG-PET images are composed of thousands of voxels (volume elements) so one is left with a situation where there are vastly more variables than subjects. Therefore, it is necessary to perform a data reduction before a statistical procedure can be applied. Approaches using both the entire image and summary statistics calculated on a number of volumes of interest (VOI) are examined. We performed the data reduction techniques of principal components analysis (PCA) and partial least-squares (PLS) on the entire image and then used linear discriminant analysis (LDA), quadratic (QDA) or logistic regression (LR) to classify subjects as having AD or FTD. Some of these methods achieve diagnostic accuracy (as assessed by leave-one-out cross-validation) that is similar to visual ratings by expert raters. Methods using PLS appear to be more successful. Averaging or using VOI data may also be helpful. Copyright © 2004 John Wiley & Sons, Ltd.

KEY WORDS: FDG-PET imaging; discriminant analysis; PLS; PCS; Alzheimer's disease

*Correspondence to: Roger Higdon, Biostatistician, National Alzheimer's Coordinating Center, 4225 Roosevelt Wy. NW, Suite 301, Seattle, WA 98105-6099, U.S.A.
†E-mail: higdonr@u.washington.edu

Contract/grant sponsor: National Alzheimer Coordinating Center; contract/grant number: AG16976
Contract/grant sponsor: Michigan Alzheimer's Disease Research Center; contract/grant number: AG08671
Contract/grant sponsor: University of California at Davis Alzheimer Center; contract/grant number: AG10129
Contract/grant sponsor: University of Pennsylvania Alzheimer Center; contract/grant number: AG10124

1. INTRODUCTION

The use of brain imaging as a diagnostic tool in neurodegenerative diseases such as Alzheimer's disease (AD) has been discussed extensively. Many studies have examined the predictive abilities of FDG-PET imaging with respect to AD and other dementing illnesses [1, 2]. Evaluation of these images is usually done through visual ratings performed by experts. Statistical classification methods have not been widely used in this area, quite possibly due to the fact images represent large amounts of data and most imaging studies have relatively few subjects (generally < 100).

The main problem is that a brain image can contain thousands of pixels (area based measurements) or voxels (volume based measurements), which represent individual measurements or variables. Therefore, the number of variables in an imaging study will generally be much larger than the number of subjects ($p \gg n$). This creates a problem since most statistical classification procedures require that the number of subjects exceed the number variables. Before a statistical classification procedure can be applied some form of data reduction must be applied to the image data. Data reduction can be successful in brain imaging because of the high degree of correlation among voxels due to both correlation of spatial structures within the brain and the filtering methods used to construct an image. Simple approaches that have been taken in the past have used either averaging or summary statistics taken over smaller brain regions. These are drawn manually based upon anatomical landmarks or by automated geometric algorithms. These approaches are referred to as region of interest (ROI) or volume of interest (VOI) analysis. In other disciplines, such as analytical chemistry [3] and DNA microarrays [4, 5] statistical methods for reducing data dimension which include principal components analysis (PCA) or partial least squares (PLS) have been combined with statistical classification methods such as linear or quadratic discriminant analysis (LDA or QDA) and logistic regression (LR). A comparison of VOI analysis with a combination of PCA and LDA has been done in an imaging study of AIDS dementia [6].

One application of imaging is the use of FDG-PET imaging to differentiate between a diagnosis of AD and a diagnosis of fronto-temporal dementia (FTD). This paper will use a multi-centre retrospective study funded by the National Alzheimer's Coordinating Center as the basis for comparing statistical methods for differentiating between AD and FTD using PET imaging.

2. OVERVIEW OF PET IMAGING

PET is a molecular imaging technique that is used to measure biochemical reactions in the body. PET imaging involves injecting a radioactive tracer into the body, then sensors that surround the body part detect positrons emitted from the tracer in opposite directions and thus can be used to localize the tracer. The PET scanner is able to quantify the intensity of emissions at each point in the body and construct an image of the body reflecting the concentration of the tracer and consequently the rate of biochemical processes. Many different tracers are used depending on what is to be measured. The most common tracer is a glucose analogue, fluorodeoxyglucose (FDG) and it is used to measure regional cerebral glucose metabolism that reflects neuronal activity.

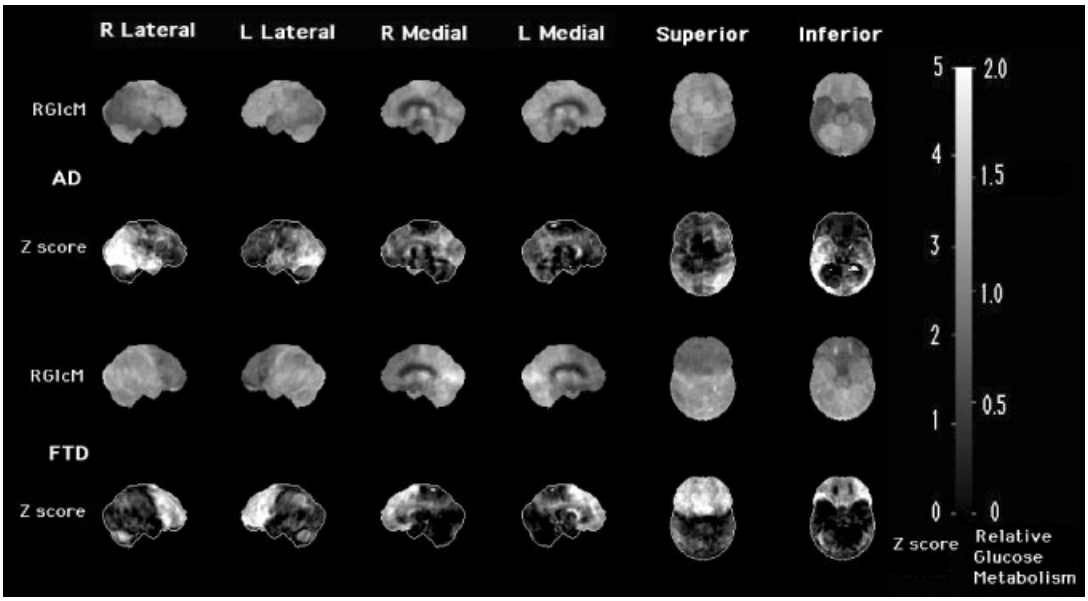

Figure 1. Examples of 3D-SSP images from a patient with Alzheimer's disease (top rows) and a patient with frontotemporal dementia (bottom rows) used in this study. Data are shown in six views of the brain on the scale as indicated on the bar at the right. The top row of images for each patient shows cerebral glucose metabolism relative to pons. The bottom row shows significant decreases in metabolism relative to control subjects as Z-scores. Note that the regions affected in these two patients differ. Physicians use the differences in these patterns to aid diagnosis.

Since each individual's brain structure varies and different PET scanners have different resolutions and fields of view, direct comparisons of PET images from different subjects can be difficult. Some form of standardization is useful for viewing PET images and necessary for conducting statistical analysis. One such approach to standardizing brain images is through linear scaling and non-linear warping [7]. This methodology stretches and contracts a brain image to match a series of landmarks inside a 'standard brain'. Grey matter activities from the standardized PET brain images are then extracted in the form of three-dimensional stereotactic surface projections (3D-SSP) [8]. This creates an image based on a standardized co-ordinate system of approximately 16 000 voxels, where voxels from two different images are directly comparable measuring metabolism at particular location within the 'standard' brain. These 3D-SSP images allow for voxel by voxel statistical comparisons. These comparisons can be redisplayed as images called statistical parametric maps. Figure 1 shows two 3D-SSP images, the first from a subject with AD and the second from a subject with FTD. The first row in each is the 3D-SSP image for six views of the brain surface (left and right lateral, left and right medial, superior and inferior). To eliminate between subject (and PET scanner) variability the 3D-SSP image is standardized by dividing by the pons value (autonomic function region) [9]. The second row gives a Z-score image, which is comprised of voxel by voxel Z-scores comparing the subject to a composite image of 33 normal subjects of similar age.

3. STATISTICAL METHODOLOGY

In order to apply statistical methods such as LDA and LR the large number of variables representing the image need to be reduced to a relative few. The reduction can be done by manual methods such as summary statistics as is done with VOI analysis or by using a variable selection procedure such as screening by between group t or F (if there are more than two classification groups) tests. Automated statistical procedures using data from the entire image such as PCA and PLS are an alternative approach. Combinations of these approaches may also be used.

3.1. Data reduction by PCA and PLS

PCA takes a linear combination of the voxels contained in the image that maximize the overall variance [10]. Successive components are found as linear transformations that are uncorrelated with the previous components. The principal components are found by taking an eigen value decomposition of either the covariance matrix or correlation matrix. In the case of imaging where the number of variables is very large the singular value decomposition is used to find eigen vectors (linear transformations).

PLS was developed for use in regression problems with large numbers of inter-correlated explanatory variables [11]. The basic PLS algorithm is as follows. Step 1: given a response variable Y and matrix of explanatory variables X (in the case of imaging the columns of X are the individual voxels), form $U_1 = Y - \bar{Y}$ and $V_1 = X - \bar{X}$ then, find a linear transformation $T_1 = V_1 w_1$ such that it maximizes the covariance between U_1 and $V_1 w_1 (w_1 = V_1' U_1 / U_1' U_1)$. Step 2: Take U_2 to be the residuals of the regression of U_1 on T_1 and let V_2 be the residuals of the separate regressions of the columns of V_1 on T_1, then repeat the process of step 1 to find w_2 and T_2. Step 3: repeat step 2 until the desired number of transformed variables is reached. In a classification problem Y is taken to be an indicator variable (0 or 1) and maximizing the covariance simplifies to maximizing the between groups sum of squares. If there are more than two groups then multiple indicator variables and a multi-variate version of the PLS algorithm can be used.

3.2. Classification by LR, LDA and QDA

The canonical variables (principal components from PCA or the T_i's from PLS) from PCA or PLS can be used as predictor variables in a classification procedure such as LDA, QDA or LR. The variables from PCA and PLS are naturally ordered, the PCA variables in terms of variance explained (size of the eigen values) and PLS by the iterations of the PLS algorithm. This creates a situation where adding variables sequentially to the LDA, QDA or LR models seems natural. Alternatively, a variable selection procedure could also be used decide which variables to add to the model.

The classification rule for LR is based on the fitted probability of classification into group 1 (assuming only two groups) or identically the log-odds of classification into group 1. Usually a subject is classified into group if the fitted probability of being in group 1 exceeds that of group 2:

$$\hat{\pi} = \frac{e^{x'\hat{\beta}}}{1 + e^{x'\beta}} > 0.5 \quad \text{or} \quad \log \frac{\hat{\pi}}{1 - \hat{\pi}} = x'\hat{\beta} > 0$$

where, π is the estimated probability of membership in group 1, x is the vector of canonical variables from PCA or PLS and β is the vector of maximum likelihood estimates from the LR model.

The classification rule for LDA is based on the linear discriminant function, which with only two groups and the assumption of normality is equivalent to $\log[f_1(x)/f_2(x)]$. Where, $f_i(x)$ is the normal density function with mean μ_i and covariance matrix Σ and x is the vector of canonical variables from PCA or PLS. Under the normality assumption the optimal classification rule is

$$\frac{f_1(x)}{f_2(x)} > \frac{1-\pi}{\pi}$$

where, π is the prior probability of membership in group 1. This can be estimated from outside data about the relative proportions of subjects from the population in groups 1 and 2 or it can be estimated from the sample proportions of subjects in groups 1 and 2.

The classification rule for that of QDA is nearly identical to LDA except that the discriminant function now has a quadratic term. This is due to the fact Σ_1 is not assumed equal to Σ_2. The classification rule then follows that of LDA above with the normal density functions not having equal covariance matrices.

When covariance matrices are not equal QDA should be a superior method to LDA. This has been shown recently in the case of DNA microarrays [5]. Under the normality assumption and equal covariance matrices LDA should be superior to LR, however if the data departs a great deal from normality then LR is usually superior. LR, LDA and QDA easily generalize to the case of more than two classification groups. References for LR and LDA (and QDA) can be found in References [12] and [10], respectively.

3.3. Assessing predictive ability

Intuitively it seems that PLS would give better predictors since it takes into account the relationship between the response and predictor variables and not just the predictor variable as in PCA when forming canonical variables. Papers in other fields of application have stated that PLS seems to perform better than PCA for classification [3, 4]. However, PLS does have serious problems with overfitting data.

To assess the predictive ability some form of cross-validation is necessary (particularly in light of the problem of PLS with overfitting data). In imaging studies since the number of subjects is small data splitting for 'true' cross-validation may not be practical. So, leave-one-out cross-validation is a reasonable alternative in these cases. Another approach for assessing predictive ability when there are only two outcomes is through the use of receiver operating characteristic (ROC) curves [13]. An ROC curve plots sensitivity (probability of a true positive test) versus one minus specificity (probability of a false positive test). The quantity area under the ROC curve can be used to assess overall predictive ability of a classification procedure (i.e. the closer to 1 the better the test). ROC curves can also be used to find 'optimal' cutoff values for the classification procedures.

3.4. Missing data

Missing data may occur in the 3D-SSP images when imaging instruments with a limited field of view are used. The simplest approach to missing data is to only include voxels with

complete data. This approach is often reasonable since the proportion of missing voxels is often low, the voxels are often from 'uninteresting' areas of the brain and the high amount of spatial correlation in images means there will be a lot of redundant information. An alternative to discarding data would to use simple mean imputation. More complicated multiple imputation procedures could be implemented which might reduce the chance of bias but because of the very large amounts of data and the complexity of numerical procedures they may not be practical.

4. THE STUDY

The study that will be used to demonstrate and assess these methods is a study using FDG-PET imaging to differentiate between a diagnosis of AD and FTD. The study is retrospective multi-centre collaborative trial based at the University of Michigan (P.I. Norman Foster, M.D.) and is funded by the National Alzheimer's Coordinating Center. The study has 48 autopsy confirmed cases (autopsy is the 'gold' standard for both the diagnosis of AD and FTD) of AD and FTD (34 AD and 14 FTD) who also have FDG-PET imaging. The study compares the accuracy of FDG-PET imaging and clinical assessment via a written descriptive scenario. Six different expert neurologists conducted classification of 3D-SSP with Z-score images of FDG-PET images by visual rating of images. The diagnostic accuracy of the six neurologists rating the 3D-SSP with Z-score images of FDG-PET images averaged 89 per cent (ranging from 87 to 91 per cent).

4.1. Description of AD and FTD

AD and FTD are both progressive neurodegenerative disorders. AD is the most common form of dementia and most commonly affects older people (65 years and up). AD impairs memory and cognitive function and is characterized pathologically by the formation of amyloid plaques and neurofibrillary tangles in the brain. FTD is pathologically diverse and includes conditions such as Pick's disease and primary progressive aphasia. FTD often occurs at younger ages than AD and predominantly affects behaviour and language. Damage is predominantly localized to frontal and anterior temporal regions in contrast to AD where the hippocampus, lateral temporal and parietal association cortex are more affected. The hippocampus is too small to be easily imaged, but the visualized contrasting patterns of hypometabolism in frontal versus more posterior regions are used diagnostically. Figure 1 shows examples of AD and FTD.

4.2. Application of statistical classification methods

Both PCA and PLS were applied to 3D-SSP image data from all 48 subjects (the training set). The PCA analysis was conducted using the covariance matrix. Whether the covariance or correlation matrix was used made relatively little difference in this case because after standardizing the 3D-SSP image by the pons value voxel variances did not vary a greatly. The first 5 canonical variables were retained for each method and fit sequentially with LR, LDA and QDA models. Models using VOI data were also compared. The VOI data was formed by averaging of voxels in 49 separate brain regions (Broadman areas) in both the left and right hemispheres for a total of 98 regions. Since this still exceeds the number of

subjects PCA and PLS were applied to the VOI data as well. Also, because the left and right regions were so highly correlated averaging of the left and right regions was examined as a method to further reduce the data.

Other aspects of data reduction and model fitting were also examined. Variable selection on the basis of between group t-statistics was applied to the PCA data. Screening of voxels on the basis of size of the between group t-statistics and spatially averaging voxels (by taking cubes of voxels of various sizes, but not on the basis of anatomy as in the VOI data) were also examined. Voxels with missing data were eliminated from the analysis leaving a total of 12 851 out of an original 15 964 in the 3D-SSP image, however, the use of mean imputation of missing voxels was also examined. Finally, two of the VOI regions were also eliminated due to missing data.

Assessment of the performance of the classification procedures was done on the basis of diagnostic accuracy using leave-one-out cross validation. This was compared to the diagnostic accuracy of predicting the training data and to the accuracy of the visual raters. A non-parametric estimate of area under the ROC curve was also calculated to compare methods. The ROC curve was calculated using leave-one-out cross-validation. The assignment of the positive account was arbitrarily chosen to be AD and the negative outcome was chosen to be FTD. The area under the curve calculations are identical regardless of the assignment.

5. RESULTS

Results are based on fitting models sequentially with up to five variables (except where noted). Results are presented giving diagnostic accuracy based on both the training set and leave-one-out cross-validation. The average diagnostic accuracy of the expert visual raters is also displayed for comparison.

Figure 2 shows the diagnostic accuracy of models fit using the entire (excluding voxels with incomplete data) 3D-SSP image. They indicated better performance by PLS than PCA particularly in combination with LDA. In that situation PLS with three cannonical variables was able to equal the accuracy of visual raters based on leave-one-out cross validation. The figure also shows how PLS overfits the training data relative to the cross-validation accuracy estimate. As seen in Figure 3, using the VOI data appears to improve accuracy somewhat for the methods using LR and there is even a slight further improvement when the left and right regions are averaged. Figure 3 also shows a simple mean imputation of missing voxels in the 3D-SSP image offers some improvement in accuracy for methods using LR. Also, one can also see that QDA and LDA perform similarly for this data set indicating that the covariance structure is not different between FTD and AD subjects.

Since the PCA variables are chosen to maximize variance and not to optimize predictive ability one might expect that adding the variables stepwise by size of between group t-statistics might be helpful. However, as is shown in Figure 4 this does not appear to be the case, particularly for the raw 3D-SSP data.

Figure 5 shows a comparison of two methods of data reduction applied prior to PLS and PCA. Screening voxels with the largest between group t-statistics does not improve diagnostic accuracy in fact it appears to be worse. Diagnostic accuracy drops off little or not at all when taking simple spatial averages, even when averaging over large numbers of voxels.

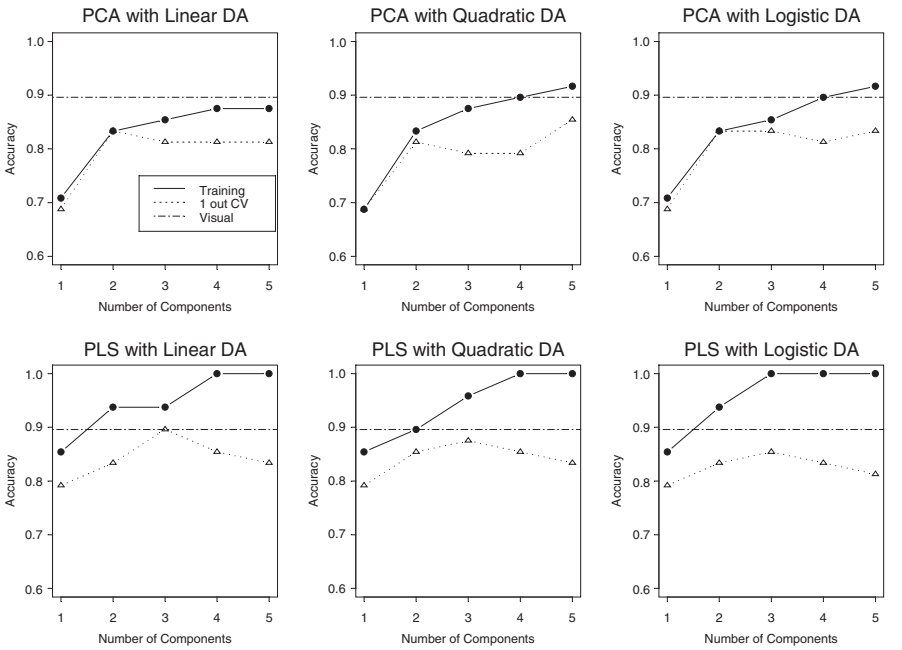

Figure 2. Plot of diagnostic accuracy using raw 3D-SSP image data (excluding voxels with incomplete data.

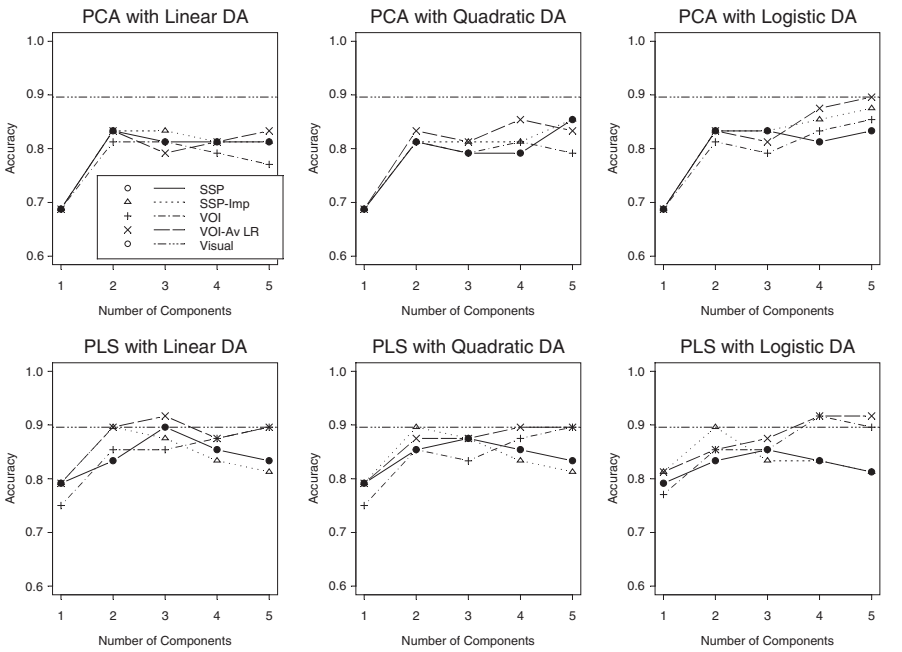

Figure 3. A plot of diagnostic accuracy based on leave-one-out cross-validation compared using the raw 3D-SSP image, 3D-SSP image with imputation, VOI data, and VOI averaging left and right.

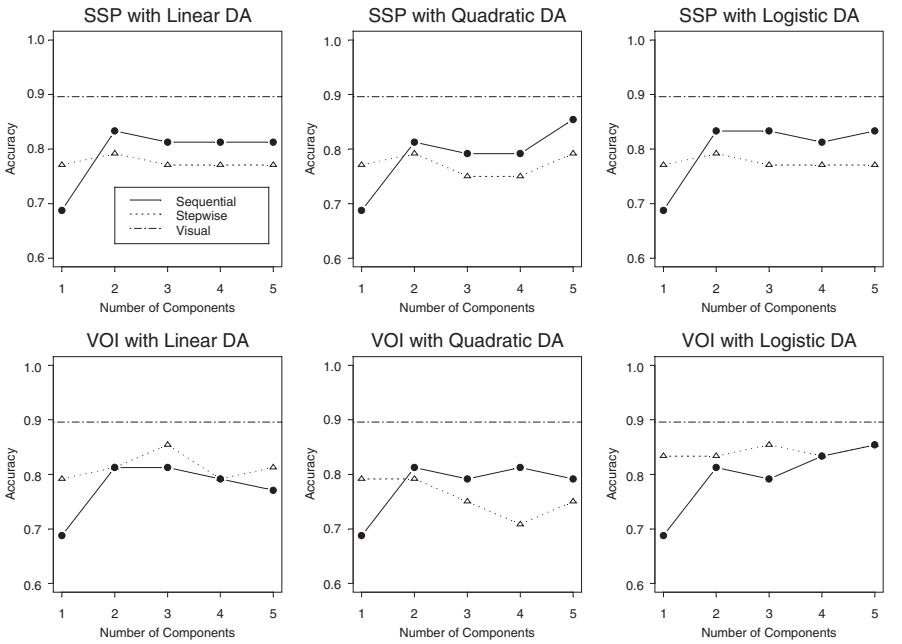

Figure 4. A plot of diagnostic accuracy comparing sequential and stepwise selection of principal components for 3D-SSP and VOI data.

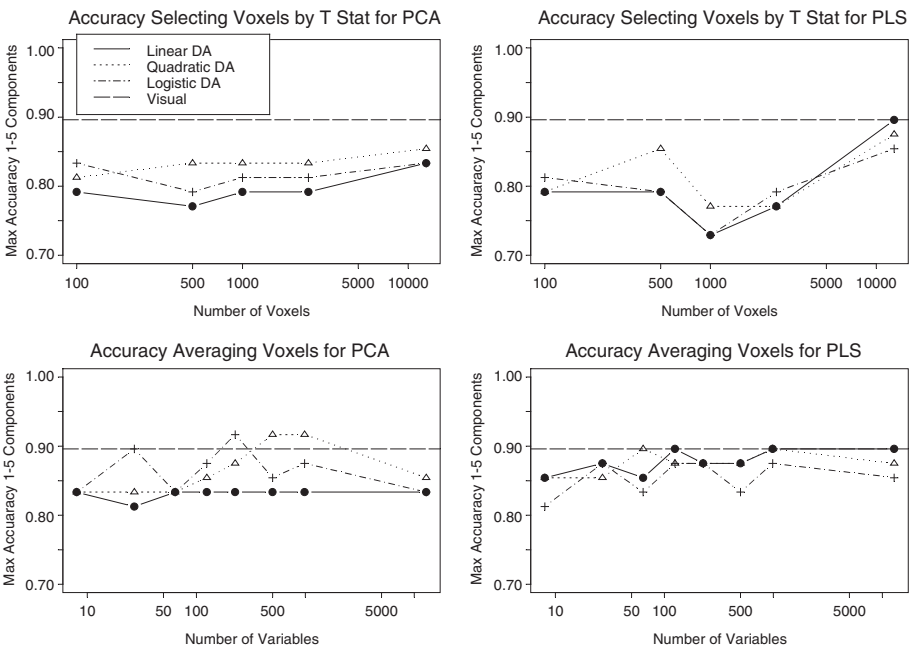

Figure 5. A comparison of the impact on accuracy of screening of voxels by *t*-test and averging of voxels plotted versus the number of voxels retained or the number of regions used.

Table I. Non-parametric AUC estimates for ROC curve calculated by leave-one-out cross-validation (maximum AUC using five components).

	3D-SSP Image	3D-SSP w/Imputation	VOI	VOI (Avg. LR)
PCA & Linear DA	0.92	0.93	0.92	0.92
PCA & Quad. DA	0.92	0.95	0.91	0.95
PCA & Logistic DA	0.92	0.91	0.90	0.90
PLS & Linear DA	0.91	0.95	0.94	0.92
PLS & Quad. DA	0.90	0.94	0.94	0.95
PLS & Logistic DA	0.91	0.90	0.94	0.90

Table II. Gain in diagnostic accuracy using 'optimal' cutoff value (note, cutoff value is chosen using the same data used to evaluate accuracy so these are likely to be overestimates).

	3D-SSP Image (per cent)	3D-SSP w/Imputation (per cent)	VOI (per cent)	VOI (Avg. LR) (per cent)
PCA & Linear DA	2	4	4	2
PCA & Quad. DA	2	4	4	4
PCA & Logistic DA	4	2	2	0
PLS & Linear DA	2	2	2	2
PLS & Quad. DA	2	2	4	2
PLS & Logistic DA	4	0	2	2

Table I gives the maximum non-parametric AUC estimates using 1–5 components for the ROC curve based on leave-one-out cross-validation. It shows that all the methods considered here are quite similar with respect to this criterion. Table II shows that some slight gains in accuracy might be achieved if an optimal cutoff is used. However, since the cutoff is chosen using the data that assesses the accuracy the estimates of gain in precision are likely too high. So, it seems little is to be gained by trying to find some sort of optimal cutoff.

6. CONCLUSIONS

The results are based only a single study so general conclusions about whether statistical classification methods are effective with imaging data or which classifications methods are best cannot be made. It is also impossible to tell whether small differences in accuracy between methods are truly indicative of a difference or merely the result of random variation. However, the data do provide some useful insight into how statistical classification methods can be applied to imaging data and show that statistical methods can give reasonably accurate diagnoses using imaging data in an applied situation.

Some previous studies have looked at the differential diagnosis of AD and FTD using imaging combined with LDA [14, 15]. These studies used single photon emission computed tomography (SPECT) for diagnosis but compared the imaging diagnosis only to the clinical diagnosis not to an autopsy diagnosis. These studies also only used VOI data and not pixel

or voxel data. These studies were able to achieve reasonably good correspondence between the clinical and model based imaging diagnoses.

It appears that methods based on PLS were slightly superior to PCA based on diagnostic accuracy whether using voxel or VOI data. The PLS methods were able to approach and in some cases equal expert visual raters even with this relatively small training set. Since, statistical methods avoid the problems of inter-rater reliability this provides some hope that statistical classification methods can be a useful tool for making diagnoses between FTD and AD or possibly other diseases using FDG-PET imaging. The weight that is applied to each voxel or VOI in the classification model may provide useful insight as to which brain regions in this case best discriminate between FTD and AD.

Images may contain a large amount of redundant information because of spatial correlation and the filtering methods used to construct images. The impact of this can be seen in how spatial averaging (VOI data) does not hinder and may improve diagnostic accuracy and how screening voxels by between group *t*-tests is so ineffective because it retains so much redundant data.

It will be important in the future to see if the diagnostic accuracy of these statistical models holds for a 'true' cross-validation set of data that is not used to construct the predictive model. Also, it may be of benefit to see if a larger more diverse training set will improve predictive abilities of the models. It remains to be seen if the methods will work well in other settings or when applied to other data sets.

It is certainly possible to try and improve or generalize these methods. Possible ways to improve these models include incorporating demographic or psychometric testing data, or by using other types of discrimination procedures such classification trees or neural networks. The methods my also be broadened to include more than two classifications, such as in this case including classification for dementia with Lewy bodies or vascular dementia. However, any increase in the amounts of data to be included or in the complexity of the models will require a substantial increase in the size of the training set required to build the models. The application of these models is not necessarily limited to classification problems certainly these procedures could be used to correlate imaging data with continuous data as well. For example, imaging data could be used predict cognitive decline or function as measured by a neuro-psychological test. As imaging is utilized more and more it will be important to develop statistical methodology to better utilize this data.

ACKNOWLEDGEMENTS

This work supported by a pilot cooperative project grant from the National Alzheimer Coordinating Center (NIH Grant AG16976) and by the Michigan Alzheimer's Disease Research Center (NIH Grant AG08671), the University of California at Davis Alzheimer Center (NIH Grant AG10129), and the University of Pennsylvania Alzheimer Center (NIH Grant AG10124). Co-author Satoshi Minoshima, MD, PhD developed and owns the copyright to the Neurostat neurological statistical imaging software (release 8/14/00) used in this project.

REFERENCES

1. Hoffman JM, Welsh-Bohmer KA, Hanson M, *et al.* FDG PET imaging in patients with pathologically verified dementia. *Journal of Nuclear Medicine* 2000; **41**(11):1920–1928.

2. Silverman DH, Small GW, Chang CY, *et al.* Positron emission tomography in evaluation of dementia: regional brain metabolism and long-term outcome. *Journal of the American Medical Association* 2001; **286**(17): 2120–2127.
3. Kemsley EK. Discriminant analysis of high-dimensional data: a comparison of principal components analysis and partial least squares data reduction methods. *Chemometrics and Intelligent Laboratory Systems* 1996; **33**(1):47–61.
4. Nguyen DV, Rocke DM. Tumor classification by partial least squares using microarray gene expression data. *Bioinformatics* 2002; **18**(1):39–50.
5. Dudoit S, Fridlyand J, Speed TP. Comparison of discrimination methods for the classification of tumors using gene expression data. *Journal of the American Statistical Association* 2002; **97**(457):77–87.
6. Liow JS, Rehm K, Strother SC, *et al*. Comparison of voxel- and volume-of-interest-based analyses in FDG PET scans of HIV positive and healthy individuals. *Journal of Nuclear Medicine* 2000;**41**(4):612–621.
7. Minoshima S, Koeppe RA, Frey KA, Kuhl DE. Anatomical standardization: linear scaling and nonlinear warping of functional brain images. *Journal of Nuclear Medicine* 1994; **35**:1528–1537.
8. Minoshima S, Frey KA, Koeppe RA, Foster NL, Kuhl DE. A diagnostic approach in Alzheimer's disease using three-dimensional stereotactic surface projections of fluorine-18-FDG PET. *Journal of Nuclear Medicine* 1995; **36**(7):1238–1248.
9. Minoshima S, Frey KA, Foster NL, Kuhl DE. Preserved pontine glucose metabolism in Alzheimer's disease: a reference region for functional brain image analysis. *Journal of Computer Assisted Tomography* 1995; **19**(4):541–547.
10. Morrison DF. *Multivariate Statistical Methods*. McGraw-Hill: New York, 1990.
11. Garthwaite PH. An interpretation of partial least squares. *Journal of the American Statistical Association* 1994; **89**(425):122–127.
12. McCullagh P, Nelder JA. *Generalized Linear Models*, 2nd edn. Chapman & Hall: London, 1989.
13. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clinical Chemist* 1993; **39**(4):561–577.
14. Pickut BA, Saerens J, Marien P, *et al*. Discriminative use of SPECT in frontal lobe-type dementia versus (senile) dementia of the Alzheimer's type. *Journal of Nuclear Medicine* 1997; **38**(6):929–934.
15. Charpentier P, Lavenu I, Defebvre L, *et al*. Alzheimer's disease and frontotemporal dementia are differentiated by discriminant analysis applied to (99m)Tc HmPAO SPECT data. *Journal of Neurology, Neurosurgery and Psychiatry* 2000; **69**(5):661–663.

