A novel events-based model for mapping disease progression and its application to familial Alzheimer's disease

H. M. Fonteijn[1], M. J. Clarkson[2], M. Modat[1], J. Barnes[2], M. Lehmann[2], S. Ourselin[1], N. C. Fox[2], and D. C. Alexander[1]
[1]Computer Science, Centre for Medical Image Computing, London, United Kingdom, [2]Institute of Neurology, Dementia Research Centre, London, United Kingdom

Introduction

In this abstract we introduce a novel method to model disease progression using cross-sectional data. Diseases typically develop through a number of stages, which can be characterized by patient symptoms and by the changes on the cellular level that cause these symptoms. An understanding of disease progression is crucial for earlier and improved diagnostic accuracy and for targeted treatment strategies. Detailed investigations of disease progression usually rely on longitudinal data that maps the entire disease progression for each patient. Cross-sectional data, on the other hand is much more convenient to collect, but only provides incomplete information about disease progression for each patient. Crude descriptions of disease progression can be obtained from cross-sectional data by classifying patients with similar symptoms into several (typically 3 to 4) disease stages and computing average measurements at each stage [1]. The key innovation of our algorithm is that it models disease progression as a series of events. We treat the data of each patient as a snapshot of this process in which some events have occurred and others have not. Then we determine the most likely series of events, given the patients' data. We demonstrate this approach on serial MRI data from a familial Alzheimer's disease (fAD) cohort, where the events are the onsets of atrophy in cortical and subcortical regions. We show progression on a much finer level than previous studies, confirming progression patterns from earlier pathological studies.

Theory

The disease progression model consists of two stages: the event detection stage and the event ordering stage. The event detection stage determines the probability that an event has occurred within each subject. The ordering algorithm constructs candidate sequences of events and uses the probability that events have occurred in each patient to calculate the likelihood of a sequence. We use a Markov Chain Monte Carlo algorithm to sample from the posterior distributions on the sequences.

In the case of the fAD data, we determine the probability of significant atrophy (which is the event of interest) by fitting a Gaussian mixture model [2] to the distribution of atrophy values in controls and patients and by calculating the probability that a data point (i.e. a regional atrophy value within a patient) is a member of the 'significant atrophy' component of the distribution. The MCMC algorithm uses 20 chains, 10^4 burn-in iterations, which we discard, and 10^5 MCMC iterations.

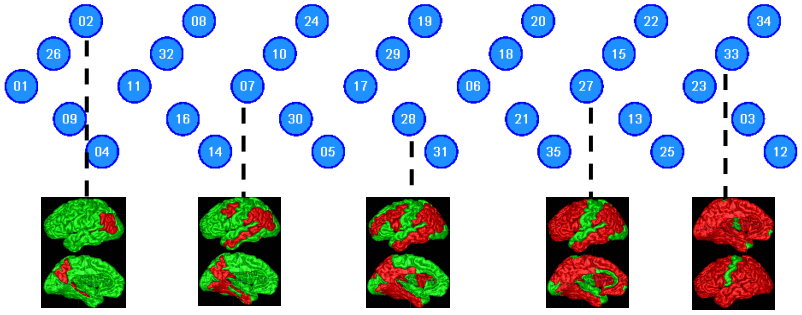
Methods

Data

The fAD data we use has been previously analyzed in Ridha et al [3]. Briefly, nine carriers of autosomal mutations associated to Alzheimer's disease were recruited together with 25 age-matched and sex-matched controls. After informed consent was given all mutation carriers underwent neuropsychological assessments, including the mini-mental state examination and volumetric MRI scans at each visit (41 visits: three to eight per patient). Each control patient had 2-4 MRI scans adding up to 54 scans in total. The images were acquired using a 1.5 T GE Signa MRI scanner using a SPGR sequence with the following parameters: (256x128 matrix, FOV=24x24 cm, TR/TE/NEX/FA=35ms/5ms/1/35°) yielding 124 contiguous 15 mm thick slices. We refer to Ridha et al. [3] for a more detailed description of this data set. Although this data set is longitudinal in nature, we treat each follow up scan as a measure of volume change from baseline and we discard the information about the temporal ordering of the follow up scans.

Preprocessing

The MRI image of each time-point is non-linearly registered to the baseline scan, using a free-form deformation method as described by [4]. We use the determinant of the Jacobian of the deformation field as a measure for expansion/contraction. We use Freesurfer to parcellate the cortex into 70 cortical regions, using anatomical landmarks and we compute the mean Jacobian for each region in each patient.



01: Hippocampus
02: bankssts
03: caudalanteriorcingulate
04: caudalmiddlefrontal
05: corpuscallosum
06: cuneus
07: entorhinal
08: fusiform
09: inferiorparietal
10: inferiortemporal
11: isthmuscingulate
12: lateraloccipital
13: lateralorbitofrontal
14: lingual
15: medialorbitofrontal
16: middletemporal
17: parahippocampal
18: paracentral
19: parsopercularis
20: parsorbitalis
21: parstriangularis
22: pericalcarine
23: postcentral
24: posteriorcingulate
25: precentral
26: precuneus
27: rostralanteriorcingulate
28: rostralmiddlefrontal
29: superiorfrontal
30: superiorparietal
31: superiortemporal
32: supramarginal
33: frontalpole
34: temporalpole
35: transversetemporal

Figure 1.Results of the events ordering algorithm. The circles indicate the order of the regions along the horizontal axis. The vertical displacement is purely to avoid overlap. Each image contains a snapshot of the atrophy pattern. Red areas are newly atrophied areas and green regions show no significant atrophy.

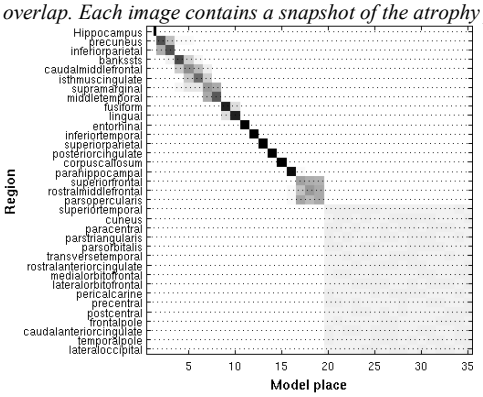


Figure 2: Histogram of regions' position in all MCMC chains

Results and Conclusions

Figure 1 shows the sequence of atrophy occurrence in fAD. The first regions to show significant atrophy are the hippocampus, the precuneus and the inferiorparietal cortex, followed by temporal regions, such as the entorhinal cortex. In the next stage of the progression pattern more and more parietal and prefrontal areas are affected. Only in the last stage are primary cortices, such as the primary motor cortex affected. This progression pattern broadly agrees with how neurofibrillary tangles (NFTs) spread through the brain as demonstrated by Braak and Braak[5]: The main difference lies in the early involvement of the precuneus and other parietal areas. Our findings are however concordant previous studies by for instance Scahill et al., who have also shown early involvement of these structures. Figure 2 shows a histogram of the position of each area in the MCMC chains, which gives an indication of the uncertainty associated with each position. This uncertainty increases in the later stages of the disease, because there are fewer data points from patients in these stages of fAD.

In conclusion, we have introduced a novel method for determining patterns of disease progression from cross-sectional data. When applying this method on structural MRI data from a fAD cohort, we show patterns of disease progression which where until now only available from post-mortem studies of pathology. Our model should therefore be very useful modeling diseases and understanding their natural history.

References: [1] Scahill R. et al. *PNAS* 2006. [2] Mclachlan P and Peel. D. Finite Mixture Models. *John Wiley and Sons.* 2000. [3] Ridha B. et al. *Lantcet Neurol* 2006 [4] Modat M. et al. *Computer Methods and Programs in Biomedicine* 2010. Braak H. & Braak E. *Acta Neuropathological* 1991