Introduction to Data-Driven Disease Progression Modelling

Neil Oxtoby
Progression Of Neurodegenerative Disease (POND) group
Centre for Medical Image Computing (CMIC)
Department of Computer Science, UCL

UCL
Centre for Medical Image Computing
Disease Progression

Aisen et al.  
Alz. Dement. 2010

Frisoni et al  
Nat. Rev. Neurol. 2010

Progression Modelling

• Construct a picture of how a disease plays out over time

• Express in terms of symptoms, pathologies and/or biomarkers

• Reconstruction ideally exploits cross-sectional data

Aisen et al. Alz. Dement. 2010
Frisoni et al. Nat. Rev. Neurol. 2010
Stitch snapshots into full picture

What we want

Disease progression

What we have

Disease progression

Time since baseline
What’s involved?

• Machine Learning, Estimation
  – Latent variable regression
  – Clustering, mixture modelling
  – Optimisation, Regularisation
  – Likelihoods
  – Bayesian statistics
Diagnosis & Staging

Patient data to match

Disease 1

Biomarker

Likelihood

Disease stage

Disease 2

Biomarker

Likelihood

Disease stage
Traditional Models

• Regress biomarker against pre-specified disease stage
  – Clinical groups: Normal / Prodromal / Symptomatic

Scanhill et al. PNAS 2002
• T1 MRI measures neuronal atrophy: subdivide using MMSE test
Traditional Models

- Regress biomarker against disease severity
  - Clinical groups: Normal / Prodromal / Symptomatic
  - Cognition:
    - MMSE (Sabuncu 2011)
    - ADAS-Cog (Yang 2011)
  - Hippocampal atrophy

Sabuncu et al. Arch. Neurol 2011
Traditional Models

• Regress biomarker against pre-specified disease stage
  – Inherited diseases: familial age of onset

Bateman et al. NEJM 2012

• Parental age of symptom onset in dominantly-inherited AD
New Models 1

- Pattern recognition: supervised learning
  - Learn to classify patients from labelled data
  - Shown value of combining imaging and non-imaging data

Classifying structural MRI in AD

Klöppel et al. Brain 2008

Disease State Fingerprint for AD

Mattila et al. JAD 2011
• Pattern discovery: unsupervised learning
  – Learn disease subtypes/stages automatically
  – Clustering

Clustering brain grey matter density to find atrophy “factors” in AD

Zhang et al. PNAS 2016
New Models 2

• Generative models
  – **Unstructured data**: scalar biomarkers, phenomenological
  – Structured data: images, connections

AD marker trajectories

Self-modelling regression

Differential Equation Models

---


Related: Jedynak et al. NeuroImage 2012

• Generative models
  – Unstructured data: scalar biomarkers, phenomenological
  – **Structured data**: spatial info. — images, brain connectivity
    • Spatiotemporal models: e.g. shape/image regression
      Durrleman et al. IJCV 2013;
      Lorenzi et al. NeuroBiol Aging 2015; etc.
    • Network propagation models: e.g. prion-like transmission, diffusion
      Raj et al. Neuron 2012;
      Iturria-Medina et al. PLOS Comp. Biol. 2014
An event-based model for disease progression and its application in familial Alzheimer's disease and Huntington's disease

Hubert M. Fonteijn a,b,c, Marc Modat a,d, Matthew J. Clarkson a,d,e, Josephine Barnes e, Manja Lehmann e, Nicola Z. Hobbs f, Rachael I. Scahill f, Sarah J. Tabrizi f,g, Sebastien Ourselfin a,d,e, Nick C. Fox e,g, Daniel C. Alexander a,b

• Estimates the order of the “events” from a cross-sectional (or short-term longitudinal) data set

Data-driven: no prior knowledge of disease stage
Event-based Model

After Fonteijn et al.
NeuroImage 2012
Event-based Model

After Fonteijn et al. NeuroImage 2012
Event-based Model

After Fonteijn et al. NeuroImage 2012
Event-based Model

Familial Alzheimer’s progression

Huntington’s progression

Fonteijn et al. NeuroImage 2012
A data-driven model of biomarker changes in sporadic Alzheimer’s disease

Alexandra L. Young,¹ Neil P. Oxtoby,¹ Pankaj Daga,¹ David M. Cash,¹,² on behalf of the Alzheimer’s Disease Neuroimaging Initiative,¹ Nick C. Fox,² Sebastien Ourselin,¹,² Jonathan M. Schott²,* and Daniel C. Alexander¹,*

1 Centre for Medical Image Computing, Department of Computer Science, University College London, Gower Street, London, WC1E 6BT, UK
2 Dementia Research Centre, UCL Institute of Neurology, University College London, 8-11 Queen Square, London, WC1N 3AR, UK
ApoE4-positive subjects: ordering supports hypothetical models
EBM: ADNI staging

Young et al. Brain 2014

Stage:
0
1-3  CSF
4-5  Rates of atrophy
6-8  Cognitive test scores
9-14 Brain volumes

Proportion

Stage

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14

CN
MCI stable
MCI prog.
AD
EBM Stage as Risk Factor

Cox Proportional Hazards Model

Stage 0
Stage 1-3 CSF
Stage 4-5 Rates of atrophy
Stage 6-8 Cognitive test scores
Stage 9-14 Brain volumes

77% accuracy
Converters vs. Non-converters
3-5 year follow up

n.oxtoby@ucl.ac.uk
Young et al. Brain 2014
Trajectories from rates of change

Oxtoby et al. Brain 2018
Differential Equation Models

Trajectories from rates of change

Oxtoby et al. Brain 2018
Trajectories from rates of change

Oxtoby et al. Brain 2018
Trajectories from rates of change

Oxtoby et al. Brain 2018
Trajectories from rates of change

Data-driven Years to Symptom Onset (log scale)

Biomarker abnormality
What we are PONDering:

- Model development
  - Discrete
  - Continuous
  - Mechanistic
  - Combinations

- Disease applications
  - Alzheimer variants (PCA, DS), Huntington’s, Parkinson’s, FTD, Prion diseases,
  - Multiple Sclerosis, Dementias, Neurodevelopment
  - Lung diseases: COPD and IPF

Healthcare

Clinical trials

Translation

Integration

Genetics  Demographics  Lifestyle  Environmental

Personalised medicine

Now and Next
What I am PONDering:

Individualised AI for Medicine

UK Research and Innovation
Future Leaders Fellowship