Multilevel Survival Analysis with Structured Penalties for Imaging Genetics data

Pascal Lu, Olivier Colliot

Multilevel Survival Analysis with Structured Penalties for Imaging Genetics data

Pascal Lu[a,b] and Olivier Colliot[a,b], for the Alzheimer's Disease Neuroimaging Initiative

[a] Sorbonne Université, Inserm, CNRS, Institut du cerveau et la moelle (ICM), AP-HP - Hôpital Pitié-Salpêtrière, Boulevard de l'hôpital, 75013, Paris, France

[b] INRIA Paris, ARAMIS project-team, 75013, Paris, France

ABSTRACT

Predicting the future occurrence of Alzheimer's disease (AD) in patients with mild cognitive impairment (MCI) is a topic of active research. Many papers have formulated this question as a classification problem: one considers a fixed time of conversion and aims to discriminate between the patients who have converted to AD at that time and those who have not. However, a clinically more relevant question is to predict the date at which a patient to AD. Survival analysis is an adequate statistical framework for such a task.

Multimodal data (imaging and genetic) provide complementary information for the prediction. While imaging data provides an estimate of the current patient's state, genetic variants can be associated to the speed of progression to AD. Although they do not provide the same type of information, most papers in classification or regression put imaging and genetic variables on the same level in order to predict the current or future patient's state.

In this work, we propose a survival model using multimodal data to estimate the conversion date to AD, by considering joint effects between the imaging and genetic modalities. We introduce an adapted penalty in the survival model, the group lasso penalty, over joint groups of genes and brain regions.

The model is evaluated on genetic (single nucleotide polymorphisms) and imaging (anatomical MRI measures) data from the ADNI database, and compared to a standard Cox model.

Keywords: Survival analysis; Cox model; Imaging genetics; Alzheimer's Disease

1. INTRODUCTION

Early diagnosis of Alzheimer's disease (AD) is an active issue in medical imaging. In Alzheimer's disease, the group of patients with mild cognitive impairment is heterogeneous, some are more likely than others to convert to AD; and therefore giving an accurate diagnosis can be difficult.[6] In this paper, we focus on predicting accurately the conversion to AD given only one time-point.

Most existing approaches to predict the conversion define the problem as a classification task at fixed time.[8–10] The problem of using classification at a fixed time is that we have to arbitrarily choose a date at which the conversion to AD is observed, we create two groups of homogeneous patients (converting before or after the fixed date), and we do not ensure that conversion in time is monotonous.

Survival models provide a regression framework which directly estimates the conversion date using only one time-point. However, by using survival models, we make several hypotheses: the conversion event is sometimes not observed (because the study ended before conversion), sometimes never occurs (some patients will never convert to AD), and only happens once. For all patients in the study, the starting point is the entry in the study, and the conversion is the time of progression of the disease to AD.

In a single visit, several kind of data can be acquired (clinical data, neuroimages, fluid biomarkers, genetic data...). Some data, such as clinical data or neuroimaging data provide a picture of the patient's state at the time they were acquired.[8] On the contrary, others data, such as genetic data, help to identify whether or not a patient could develop AD in the future (for instance, some alleles of APOE increase the risk of developing

Further author information:
Pascal Lu: E-mail: pascal.lu@outlook.com (correspondence)
Olivier Colliot: E-mail: olivier.colliot@upmc.fr

AD). An issue raised by collecting data from different sources concerns their combination. In the area of imaging genetics, most papers focus either on the association between neuroimaging covariates and genetic data,[7,11] or on building machine learning predictors for a disease at fixed time using classification (logistic regression, SVM[12]). All these models combine neuroimaging and genetic data by using an additive framework. However, adding the effect of both modalities and putting them on the same level is not optimal, as these modalities do not provide the same type of information. We proposed a multilevel framework[14] for combining imaging and genetic data for classification.

In this paper, we propose a survival model, based on the Cox Proportional Hazard model and using a multilevel framework. Survival models (Cox Proportional Hazard model) have been applied for combining multimodal data[16] and for predicting the conversion to AD;[15] in both case using an additive framework. Learning the conversion date to AD with modalities taken separately shows that the genetic modality has weaker predictive value than the neuroimaging modality. Instead of summing both genetic and neuroimaging contributions (which could lead to a weaker contribution of the genetic modality in the model), we propose that the parameters, combined with the neuroimaging covariates, could be modulated by the patient's genetic data. This hypothesis leads to a multilevel model where genetic data express themselves through interactions with neuroimaging covariates.

Adding interactions leads to high-dimensional models, and adapted penalties for each modality is essential to avoid overfitting. For instance, SNPs can be grouped by genes[13,14] and a group lasso penalty can be applied on the groups formed by genes. In this paper, we will use the group lasso penalty on the interactions for the parameters coupling (genes, brain region). We use a proximal gradient descent algorithm to learn all the parameters. This model is evaluated on genetic single-nucleotide polymorphisms (SNPs) and neuroimaging data (MRI modality) from the ADNI database (`http://adni.loni.usc.edu`), and is compared to standard Cox models.

2. METHODS

2.1 Model set-up

We aim to model the time to conversion to Alzheimer's Disease (AD) for MCI patients. For all patients, the starting point is the entry in the study, and the conversion is the time of progression to AD.

Notations For each patient i, we denote T_i^* his real conversion date from MCI to AD, C_i the date of his final visit and $T_i = \min(T_i^*, C_i)$ the duration observed in the study. We introduce $\delta_i = \mathbb{I}\{T_i^* \leq C_i\}$ indicating if the conversion has occurred.

We denote $\mathbf{x}_{\mathcal{G}}^i$ the vector of single-polymorphism nucleotides (SNP) counted by number of minor variants, $\mathbf{x}_{\mathcal{I}}^i$ the vector of imaging variables (brain regions), $|\mathcal{G}|$ the number of SNPs and $|\mathcal{I}|$ the number of imaging variables.

The conversion date T is a continuous random variable with cumulative distribution function $F : t \mapsto \mathbb{P}\{T < t|\mathbf{x}_{\mathcal{G}}^i, \mathbf{x}_{\mathcal{I}}^i\} = 1 - \exp\left(-\int_0^t h(u|\mathbf{x}_{\mathcal{G}}^i, \mathbf{x}_{\mathcal{I}}^i)\mathrm{d}u\right)$ where h is the hazard function, representing the instantaneous rate of occurrence of the event.

Multilevel framework We propose the multilevel framework, based on the Cox proportional hazard assumption, defined by $h(t|\mathbf{x}_{\mathcal{G}}^i, \mathbf{x}_{\mathcal{I}}^i) = h_0(t)\mathrm{e}^{\boldsymbol{\beta}(\mathbf{x}_{\mathcal{G}}^i)^\top \mathbf{x}_{\mathcal{I}}^i}$, where $\boldsymbol{\beta}(\mathbf{x}_{\mathcal{G}}^i)$ is the parameter vector depending on genetic data $\mathbf{x}_{\mathcal{G}}^i$ and h_0 is the baseline hazard function describing the risks for individuals whose covariates are null.

We make the assumption that $\boldsymbol{\beta}$ is an affine function depending on genetic data: $\boldsymbol{\beta}(\mathbf{x}_{\mathcal{G}}^i) = \mathbf{W}^\top \mathbf{x}_{\mathcal{G}}^i + \boldsymbol{\beta}_{\mathcal{I}}$, where $\mathbf{W} \in \mathcal{M}_{|\mathcal{G}|,|\mathcal{I}|}(\mathbb{R})$. Then,

$$h(t|\mathbf{x}_{\mathcal{G}}^i, \mathbf{x}_{\mathcal{I}}^i) = h_0(t)\mathrm{e}^{(\mathbf{x}_{\mathcal{G}}^i)^\top \mathbf{W}\mathbf{x}_{\mathcal{I}}^i + \boldsymbol{\beta}_{\mathcal{I}}^\top \mathbf{x}_{\mathcal{I}}^i}$$

The survival function is given by:

$$S(t|\mathbf{x}_{\mathcal{G}}^i, \mathbf{x}_{\mathcal{I}}^i) = (S_0(t))^{(\mathbf{x}_{\mathcal{G}}^i)^\top \mathbf{W}\mathbf{x}_{\mathcal{I}}^i + \boldsymbol{\beta}_{\mathcal{I}}^\top \mathbf{x}_{\mathcal{I}}^i} \text{ where } S_0(t) = \exp\left(-\int_0^t h_0(u)du\right)$$

The effect of the covariate $(\mathbf{x}^i_{\mathcal{G}}, \mathbf{x}^i_{\mathcal{I}})$ on the survival function is to raise it to a power given by the prognostic index $\mathrm{PI}(\mathbf{x}^i_{\mathcal{G}}, \mathbf{x}^i_{\mathcal{I}}) = \mathrm{e}^{(\mathbf{x}^i_{\mathcal{G}})^\top \mathbf{W} \mathbf{x}^i_{\mathcal{I}} + \boldsymbol{\beta}_{\mathcal{I}}^\top \mathbf{x}^i_{\mathcal{I}}}$.

The genetic modality has a much weaker predictive power compared to imaging or clinical features. And separately taken, the genetic modality provides poor results (see table 1). By combining SNPs and imaging features in that way, we ensure that SNPs will add a significant contribution to the model.

2.2 Optimization

Given the dataset $\{(\mathbf{x}^i_{\mathcal{G}}, \mathbf{x}^i_{\mathcal{I}}, T_i, \delta_i), i = 1, \ldots, N\}$ where the covariates $\mathbf{x}_{\mathcal{G}}, \mathbf{x}_{\mathcal{I}}$ and $\mathbf{x}_{\mathcal{G}} \mathbf{x}_{\mathcal{I}}^\top$ are centered and normalized, the negative partial log-likelihood is given by:

$$\ell(\mathbf{W}, \boldsymbol{\beta}_{\mathcal{I}}) = -\frac{1}{N} \sum_{i=1}^{N} \delta_i \left((\mathbf{x}^i_{\mathcal{G}})^\top \mathbf{W} \mathbf{x}^i_{\mathcal{I}} + \boldsymbol{\beta}_{\mathcal{I}}^\top \mathbf{x}^i_{\mathcal{I}} - \log \sum_{j \in \mathcal{R}(T_i)} \mathrm{e}^{(\mathbf{x}^i_{\mathcal{G}})^\top \mathbf{W} \mathbf{x}^i_{\mathcal{I}} + \boldsymbol{\beta}_{\mathcal{I}}^\top \mathbf{x}^i_{\mathcal{I}}} \right)$$

where $\mathcal{R}(T_i)$ is the set of patients j such that $T_j \geq T_i$.

Penalties As the number of parameters to estimate is much larger than the number of patients, we need to add penalties on $\mathbf{W}$ and $\boldsymbol{\beta}_{\mathcal{I}}$. For imaging parameters $\boldsymbol{\beta}_{\mathcal{I}}$, we considered the ridge penalty, as Alzheimer's Disease has a diffuse anatomical pattern of alteration. For the matrix $\mathbf{W}$, we start by mapping SNPs to genes $\mathcal{G}_\ell$ ($\ell \leq L$, where L is the number of genes), and we use a group lasso with overlap penalty, where groups are (genes, imaging covariate). This penalty enforces sparsity between groups and regularity inside the same group. Finally, we add the following penalty to the negative partial log-likelihood:

$$\Omega(\mathbf{W}, \boldsymbol{\beta}_{\mathcal{I}}) = \lambda \sum_{i=1}^{|\mathcal{I}|} \sum_{\ell=1}^{L} \sqrt{|\mathcal{G}_\ell|} \|\mathbf{W}_{\mathcal{G}_\ell, i}\|_{\ell_2} + \lambda_{\mathcal{I}} \|\boldsymbol{\beta}_{\mathcal{I}}\|_{\ell_2}^2$$

where $\lambda > 0, \lambda_{\mathcal{I}} > 0$ are the hyperparameters.

The parameters $\mathbf{W}, \boldsymbol{\beta}_{\mathcal{I}}$ are obtained by minimizing the quantity $\ell(\mathbf{W}, \boldsymbol{\beta}_{\mathcal{I}}) + \Omega(\mathbf{W}, \boldsymbol{\beta}_{\mathcal{I}})$. The usual approach for dealing with the penalty Ω is to use a proximal gradient descent on the convex set defined by Ω.[1,2]

Algorithm 1: Optimization procedure

Input: $\{(\mathbf{x}^i_{\mathcal{G}}, \mathbf{x}^i_{\mathcal{I}}, T_i, \delta_i), i = 1, \ldots, N\}$, $\delta = 0.5$, $\varepsilon_0 = 0.01$, $\eta = 10^{-5}$;
Initialization: $\mathbf{W} = \mathbf{0}$, $\boldsymbol{\beta}_{\mathcal{I}} = \mathbf{0}$, converged = False
while not(converged) **do**
  $\boldsymbol{\gamma} = (\texttt{flatten}(\mathbf{W}), \boldsymbol{\beta}_{\mathcal{I}})$ and $\boldsymbol{\omega} = \boldsymbol{\gamma} - \varepsilon \nabla \ell$;
  $\widehat{\mathbf{W}}_{\mathcal{G}_\ell, i} = \max\left(0, 1 - \frac{\varepsilon \lambda_{\mathcal{G}} \theta_{\mathcal{G}_\ell}}{\left\| \boldsymbol{\omega}^{(t)}_{\mathcal{G}_\ell + i|\mathcal{G}|} \right\|_2}\right) \boldsymbol{\omega}_{\mathcal{G}_\ell + i|\mathcal{G}|}$ for $(i, \ell) \in [\![1, |\mathcal{I}|]\!] \times [\![1, L]\!]$;
  $\widehat{\boldsymbol{\beta}}_{\mathcal{I}} = \frac{\boldsymbol{\omega}_{\mathcal{I} + |\mathcal{G}||\mathcal{I}|}}{1 + 2\varepsilon\lambda_{\mathcal{I}}}$ (imaging modality) ;
  if $(\ell + \Omega)(\widehat{\mathbf{W}}, \widehat{\boldsymbol{\beta}}_{\mathcal{I}}) > (\ell + \Omega)(\mathbf{W}, \boldsymbol{\beta}_{\mathcal{I}})$ **then**
    $\varepsilon = \delta\varepsilon$
  else
    $\mathbf{W} = \widehat{\mathbf{W}}$, $\boldsymbol{\beta}_{\mathcal{I}} = \widehat{\boldsymbol{\beta}}_{\mathcal{I}}$, $\varepsilon = \varepsilon_0$;
  end
end
converged = $\left| (\ell + \Omega)(\mathbf{W}, \boldsymbol{\beta}_{\mathcal{I}}) - (\ell + \Omega)(\widehat{\mathbf{W}}, \widehat{\boldsymbol{\beta}}_{\mathcal{I}}) \right| \overset{?}{<} \eta |(\ell + \Omega)(\mathbf{W}, \boldsymbol{\beta}_{\mathcal{I}})|$

Implementation We flatten the cross-product covariates $\mathbf{x}_{\mathcal{G}}\mathbf{x}_{\mathcal{I}}^{\top}$ and the matrix $\mathbf{W}$ and transform them into a vector. We create a vector $\boldsymbol{\gamma} = (\texttt{flatten}(\mathbf{W}), \boldsymbol{\beta}_{\mathcal{I}})$ containing the coefficients of $\mathbf{W}, \boldsymbol{\beta}_{\mathcal{I}}$. To recreate $\mathbf{W}$, we just need to unflatten $\boldsymbol{\gamma}$. The vector $\boldsymbol{\gamma}$ is updated using a proximal gradient descent described in Algorithm 1. The stopping criterion for this algorithm is

$$\left|(\ell + \Omega)(\mathbf{W}, \boldsymbol{\beta}_{\mathcal{I}}) - (\ell + \Omega)(\widehat{\mathbf{W}}, \widehat{\boldsymbol{\beta}}_{\mathcal{I}})\right| < \eta|(\ell + \Omega)(\mathbf{W}, \boldsymbol{\beta}_{\mathcal{I}})|$$

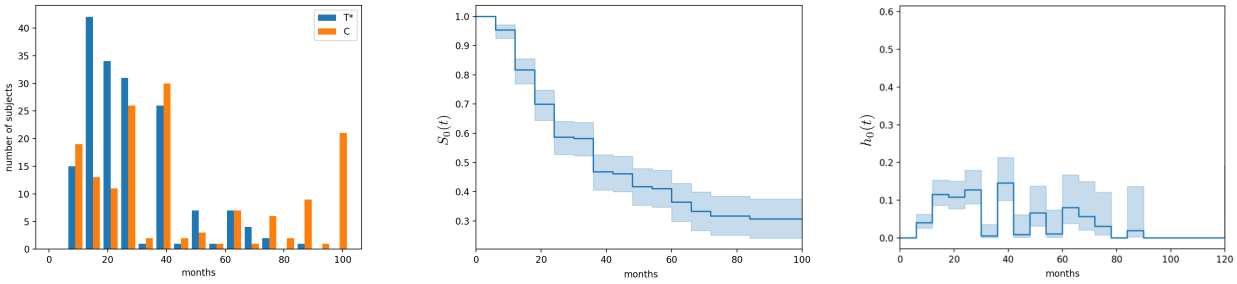
3. EXPERIMENTS AND RESULTS

3.1 Dataset

The ADNI1 GWAS dataset contains 326 MCI patients with 172 MCI patients at baseline who progressed to AD and 154 MCI patients who remained stable (censored data).

Covariates In this dataset, 620,901 SNPs have been genotyped. Based on the 44 first top genes related to AD (from AlzGene, http://www.alzgene.org) and on the Illumina annotation using the Genome build 36.2, we select 1,107 SNPs. For cross-validation purposes, SNPs for whose the variance among the dataset is smaller than 0.01 are removed, leading to 679 SNPs. Regarding the MRI modality, we use the segmentation of FreeSurfer which gives the volume of subcortical regions (44 features) and the average thickness in cortical regions (68 features).

Baseline survival function S_0 The baseline survival function S_0 is computed using the Kaplan-Meier estimate. On figure 1, is shown on the right the Kaplan-Meier estimated baseline survival function S_0 using the distribution of T^* and C displayed on the left. The follow-up date is $\tau_{\text{hor}} = 100$ months; patients who convert after this date are truncated. The median survival time (the smallest survival time for which the survivor function is less than or equal to 0.5) is 36 months.



(a) Histogram of T^* and C (b) Kaplan-Meier estimated baseline survival function S_0 (c) Nelson-Aalen estimated hazard function h_0 (bandwidth = 6 months)

Figure 1: ADNI1 Dataset: baseline survival function and hasard function

3.2 Evaluation

Baseline models We compare the multilevel framework to the Cox Proportional Hazard model using one modality or using an additive framework. In this later case, the hazard function for patient i is given by $h(t|\mathbf{x}_{\mathcal{G}}^i, \mathbf{x}_{\mathcal{I}}^i) = h_0(t)\mathrm{e}^{\boldsymbol{\beta}_{\mathcal{G}}^{\top}\mathbf{x}_{\mathcal{G}}^i + \boldsymbol{\beta}_{\mathcal{I}}^{\top}\mathbf{x}_{\mathcal{I}}^i}$.

Metrics We define the three following measures to asses the quality of the prediction:

- the concordance index (or C-index)[5] which checks if the model orders the conversion dates in the same order as the ground truth. As a generalization of AUC, the range of the C-index is $[0, 1]$, but typical values are between 0.55 and 0.7.

- the integrated Brier score,[4] for uncensored data:

$$\text{Brier} = \frac{1}{\tau_{\text{hor}}} \int_0^{\tau_{\text{hor}}} \left[\frac{1}{N} \sum_{i=1}^{N} \delta_i (\mathbf{1}(T_i^* > t) - \widehat{S}(t|\mathbf{x}_{\mathcal{G}}, \mathbf{x}_{\mathcal{I}}))^2 \right] \mathrm{d}t$$

  The Brier score measures the accuracy of probabilistic predictions. The model performs better when the Brier score is lower.

- the integrated Area Under Curve, defined as $\text{iAUC} = \int_0^{\tau_{\text{hor}}} \text{AUC}(t) f(t) \mathrm{d}t$ where τ_{hor} is the follow-up date, f is the probability density function of T and $\text{AUC}(t)$ is the cumulative AUC.[3] The cumulative/dynamic AUC plays the same role as the classical Area Under Curve in classification. As for the AUC, the range of the iAUC is $[0, 1]$, and the higher the iAUC is, the more predictive the model is.

Cross validation To determine the hyperparameters, we use a nested cross validation. We perform a 5-fold cross validation, and within each fold, we find the optimal hyperparameters using a 5-fold cross validation on the training set and taking the hyperparameters that maximize the C-index over the inner test set.

The hyperparameters are optimized between $\{10^{-4}, 10^{-3}, \ldots, 10, 10^2\}$.

Modality	Method	C-index	Brier score	iAUC
SNPs only	Cox PH model (ℓ_1 penalty)	0.521 ±0.040	0.166 ±0.009	0.515 ±0.031
MRI only	Cox PH model (no penalty)	0.636 ±0.034	0.190 ± 0.017	0.636 ±0.050
MRI only	Cox PH model (ℓ_2 penalty)	0.671 ± 0.022	0.149 ±0.008	0.663 ±0.044
All	Additive Cox PH model (ℓ_1 penalty)	0.677 ± 0.020	0.148 ± 0.006	0.680 ±0.030
All	Multilevel model (ours)	0.681 ±0.018	0.147 ± 0.006	0.686 ±0.031

Table 1: Results for different modalities and methods (mean value across the test folds ± standard deviation)

Results on table 1 show that genetic modality, taken alone, have a much weaker predictive value than the imaging modality. The imaging modality already provides good performances, and adding the genetic modality improves the model performance, but not significantly (in both additive and multilevel frameworks). Adding a penalty on the imaging parameter also increases the performances. Finally, the multilevel model provides slightly better results than the additive model.

3.3 Effect of cross-product covariates on conversion

For interpretation purposes, we compute a new reduced matrix $\widetilde{\mathbf{W}} \in \mathcal{M}_{|\mathcal{I}|,L}(\mathbb{R})$, shown on figure 2, where for each brain region j and gene ℓ, $\widetilde{\mathbf{W}}_{j,\ell} = \max_{s \in \mathcal{G}_\ell} |\mathbf{W}_{s,j}|$.

The matrix on figure 2 shows that some rows and columns have more non-null coefficients than others.

The strongest effects are found for the ventricles, which are enlarged in AD but also in aging and other degenerative diseases, and several medial temporal lobes (MTL) structures (entorhinal cortex, hippocampus, amygdala) which are altered early in AD. It is interesting to note that the ventricles have a strong interaction with all genes except APOE, while the MTL structures have interactions with a more restricted but quite consistent set of genes. Regarding the genes, the strongest effects are found for the single-nucleotide polymorphism rs6503018 (TNK1), rs429358 (APOE) and rs3093662 (TNF).

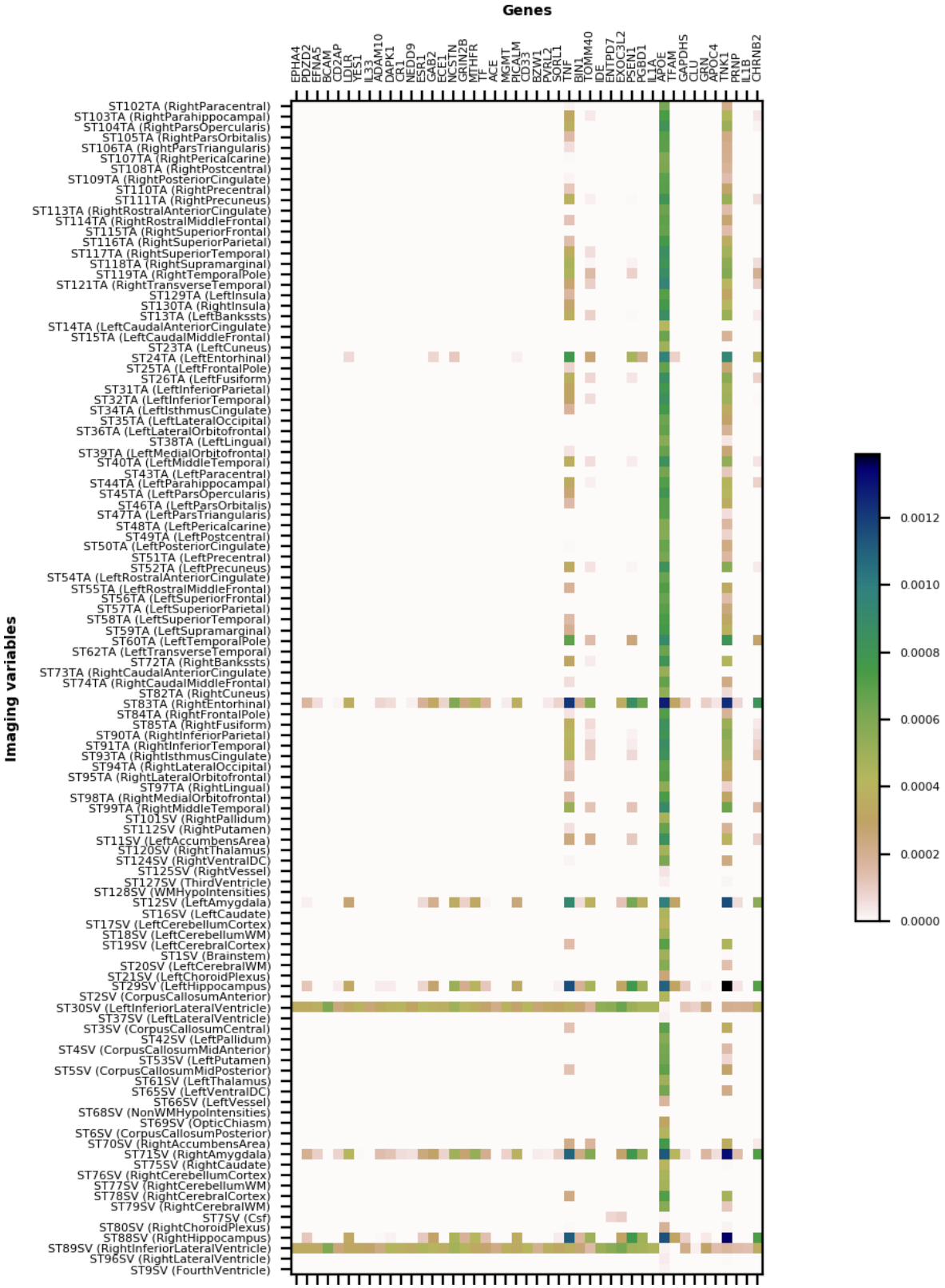


Figure 2: Reduced matrix $\widetilde{\mathbf{W}}$

4. CONCLUSION

This paper proposes a novel approach to estimate the conversion date to AD for MCI patients, using the Cox Proportional Hazard model, from genetic and neuroimaging data. On the contrary of additive models, the multilevel model captures interactions between genes and brain regions. The use of adapted penalties avoids overfitting by providing a sparse matrix and highlighting brain regions and genes both related to the progression to AD.

REFERENCES

1. Hastie, T., Tibshirani, R., Wainwright, M.: Statistical Learning with Sparsity - The Lasso and Generalizations, vol. 143. CRC Press, Boca Rato (2015)
2. Beck, A., Teboulle, M.: Gradient-based algorithms with applications to signal recovery problems. In: Palomar, D.P., Eldar, Y.C. (eds.) Convex Optimization in Signal Processing and Communications, pp. 42–88. Cambribge University Press, Cambridge (2010)
3. Chambless, L. E. et al: Estimation of time-dependent area under the ROC curve for long-term risk prediction. *Statistics in Medicine* (2006)
4. Graf E., Assessment and Comparison of Prognostic classification schemes for survival data, *Statistics in Medicine*, **18**:2529–2545 (1999)
5. Steck, H. et al: On Ranking in Survival Analysis: Bounds on the Concordance Index. *Advances in Neural Information Processing Systems* **20** (2007)
6. Razvan V. et al, TADPOLE Challenge: Prediction of Longitudinal Evolution in Alzheimer's Disease, arXiv:1805.03909 (2018)
7. Liu, J., Calhoun, V.D.: A review of multivariate analyses in imaging genetics. Front. Neuroinform. **8**(29), 1–11 (2014)
8. Rathore, S et al: A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages. Neuroimage, **155**, 530–548 (2017)
9. Cheng, B. et al. Domain transfer learning for MCI conversion prediction. *IEEE Trans. Biomed. Eng.* 62, 1805–1817 (2015)
10. Davatzikos, C., et al. Prediction of MCI to AD conversion, via MRI, CSF biomarkers, and pattern classification. *Neurobiol. Aging* 32 (2322), 2319–2327 (2011)
11. Batmanghelich, N.K., Dalca, A., Quon, G., Sabuncu, M., Golland, P.: Probabilistic modeling of imaging, genetics and diagnosis. IEEE TMI **35**, 1765–1779 (2016)
12. J. Peng, et al. Structured sparse kernel learning for imaging genetics based Alzheimers Disease Diagnosis. In: Ourselin, S., et al. (eds.) MICCAI 2016. LNCS (9901) 70–78. Springer (2016)
13. Silver, M., Janousova, E., Hua, X., Thompson, P.M., Montana, G., and ADNI: Identification of gene pathways implicated in Alzheimer's disease using longitudinal imaging phenotypes with sparse regression. NeuroImage **63**(3), 1681–1694 (2012)
14. Lu, P. et al.: Multilevel Modeling with Structured Penalties for Classification from Imaging Genetics Data. In: Cardoso M. et al. (eds) *Graphs in Biomedical Image Analysis, Computational Anatomy and Imaging Genetics*. Lecture Notes in Computer Science. **10551**:230–239 (2017)
15. Li K. et al: A prognostic model of AD relying on multiple longitudinal measures and time-to-time event data. *Alzheimer's & Dementia* **14**: 644–651 (2018)
16. Bovelstad, H. M. et al: Survival prediction from clinico-genomic models - a comparative study. *BMC Bioinformatics*. **10**(413) (2009)