

マルチブロック成分法による 脳画像解析

佐賀大学医学部

地域医療科学教育研究センター

川口淳

2018年度統計関連学会連合大会

2018年9月13日

中央大学・後楽園キャンパス

内容

- 複数のデータセットと統合解析
 - » 脳画像解析
 - > 脳の構造, 機能, 脳腫瘍解析

- マルチブロック成分分析
 - » 主成分分析の拡張(アルゴリズム)
 - » 脳画像データへの応用例

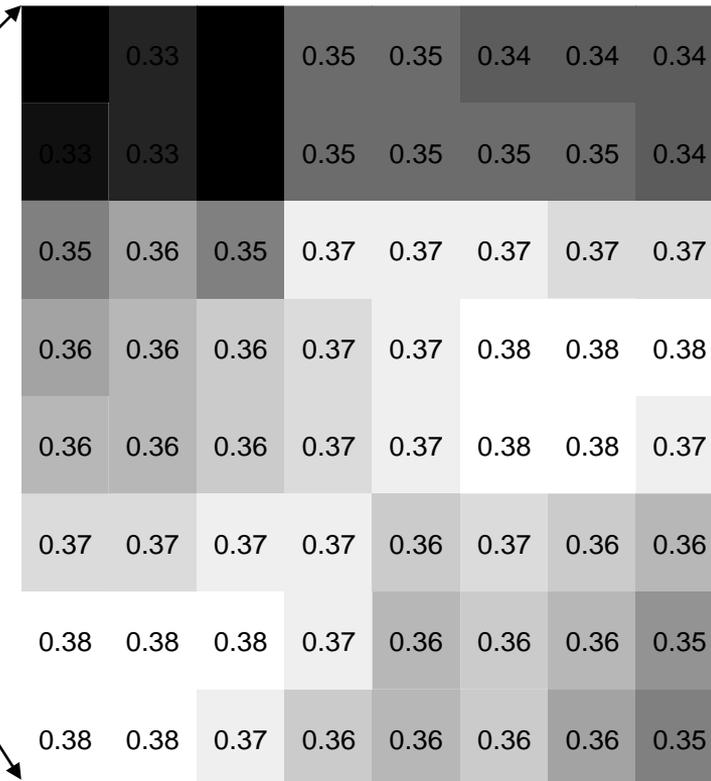
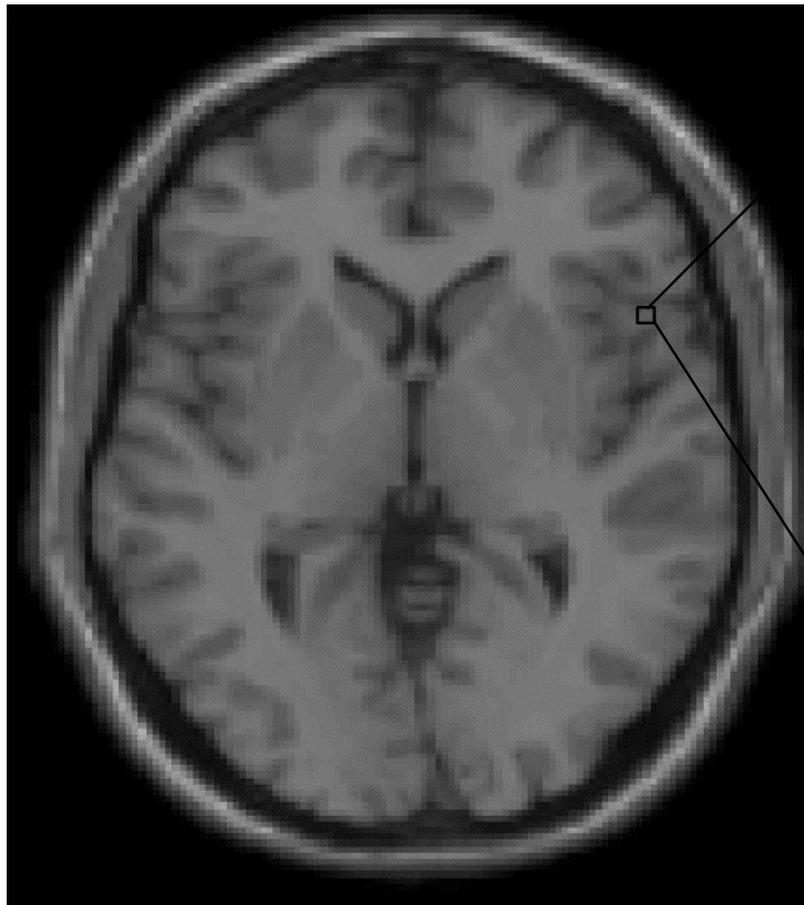
脳画像データ

川口淳 (2018). 画像データの解析.
医学統計学ハンドブック; 714-726

(B) 横断面(1スライス)

(C) 色と数値を対応

縦



横

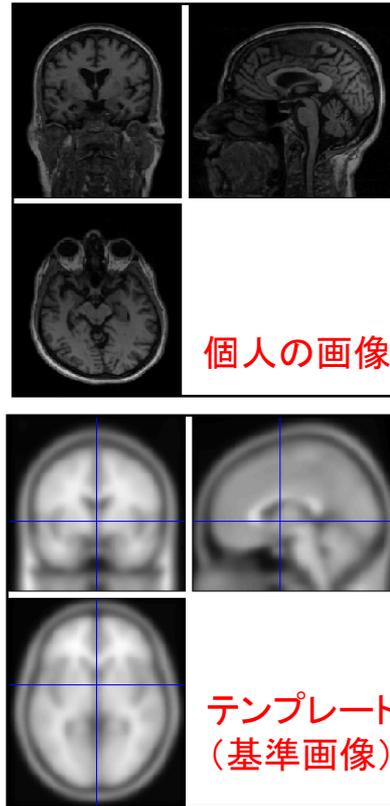
脳画像統計解析の全体図

川口(2012 計量生物学)

撮像

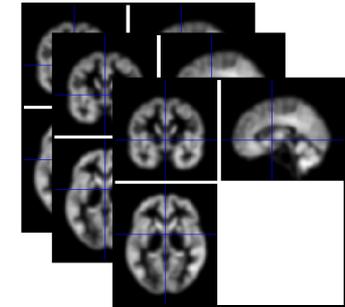


前処理

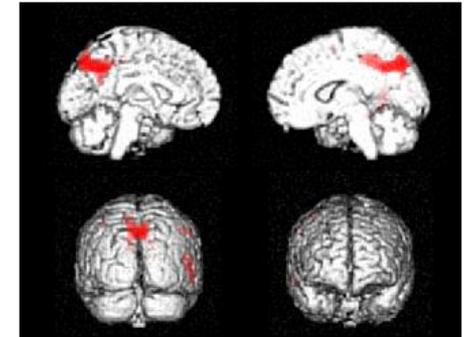


本解析

複数の対象者



解析

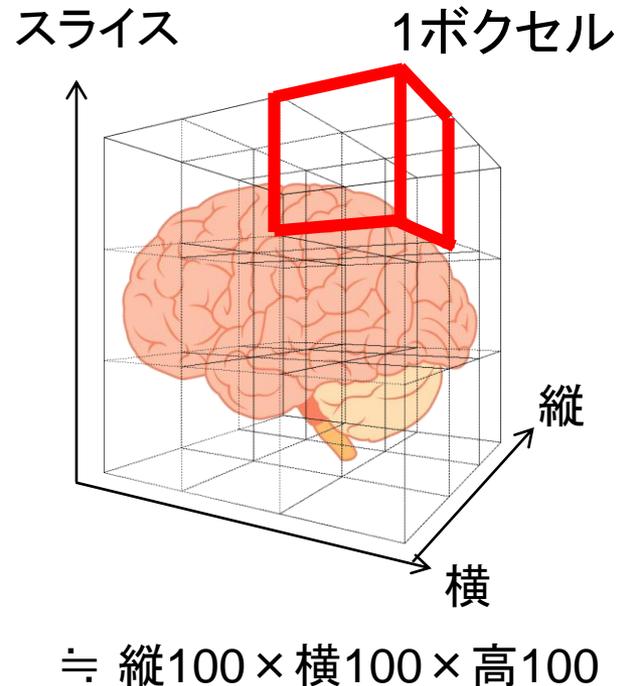


結果の表示

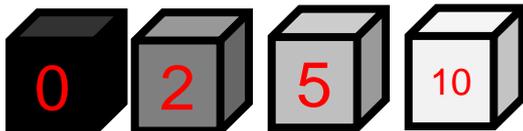
- テンプレートに色づけ

解剖学的標準化

- ボクセルの位置合わせ
- 個人間の違いを補正



ミリオンボクセル



マルチモダル脳画像解析

➤ 画像の種類

- » (形態, 構造) sMRI, dMRI (機能) fMRI, MEG, FDG/Amyloid/Tau PET

➤ 適用疾患

- » 精神障害 (認知症, 統合失調症, 気分障害, 物質関連障害)
- » 発達障害 (注意欠陥・多動性障害 (ADHD), 自閉症スペクトラム (ASD))
- » 不安障害, 境界性パーソナリティ障害

➤ 解析方法

- » SVM, Multi-kernel SVM, Random Forest, deep learning,
- » PCA, CCA, PLS, ICA, NMF...

Arbabshirani et al. (2017 NeuroImage), Meng et al. (2017 NeuroImage)
Calhoun and Sui (2016 Bio. Psy.) Vandenberghe Marsden (2015 Ph. Med. Biol.)

例① Alzheimer早期発見

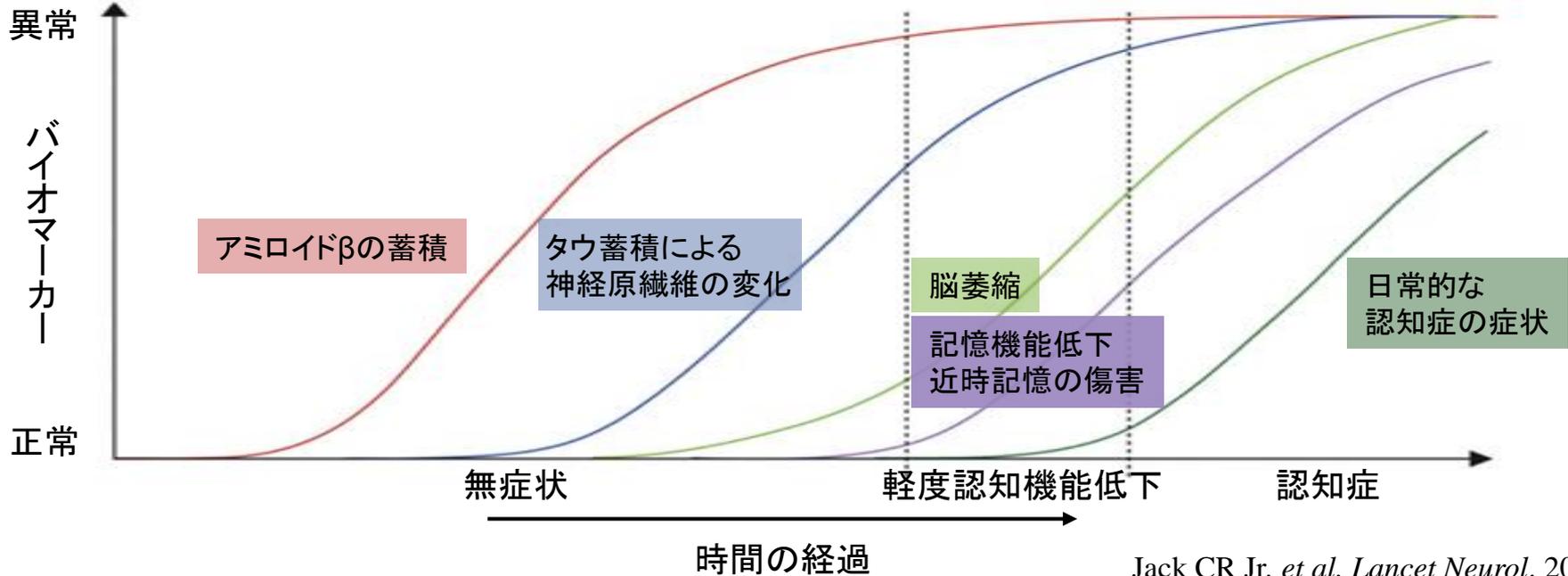
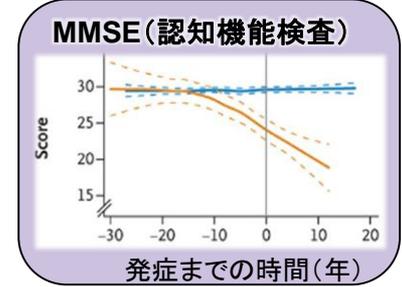
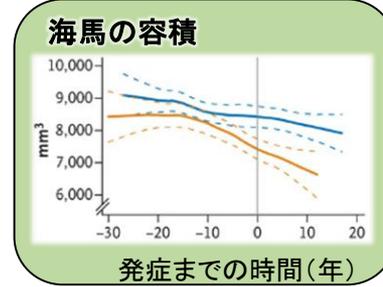
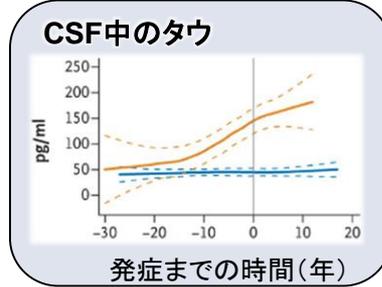
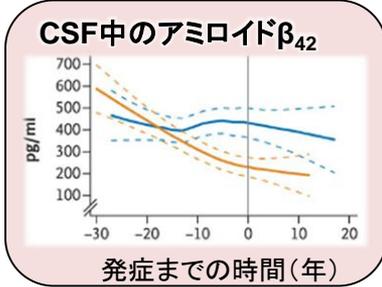
診断方法

病理診断(死後, 生前嗅粘膜生検)
PIB-PET・CSF採取

MRI

認知機能検査など

リスク遺伝子*の有無
なし — 青線
あり — 赤線



Jack CR Jr, et al. Lancet Neurol. 2010

Bateman, et al. N Engl J Med. 2012

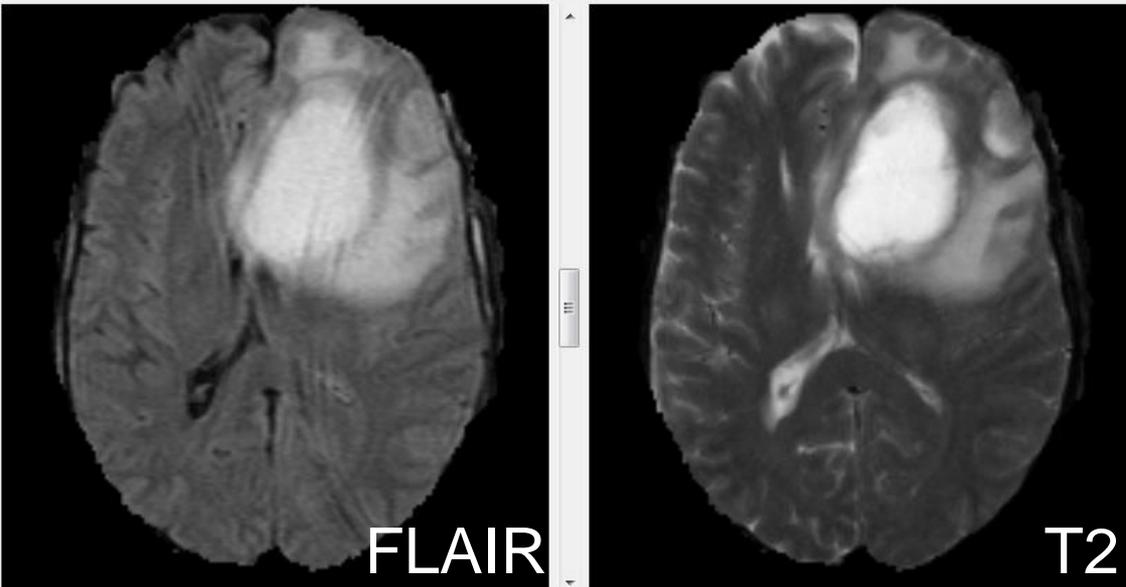
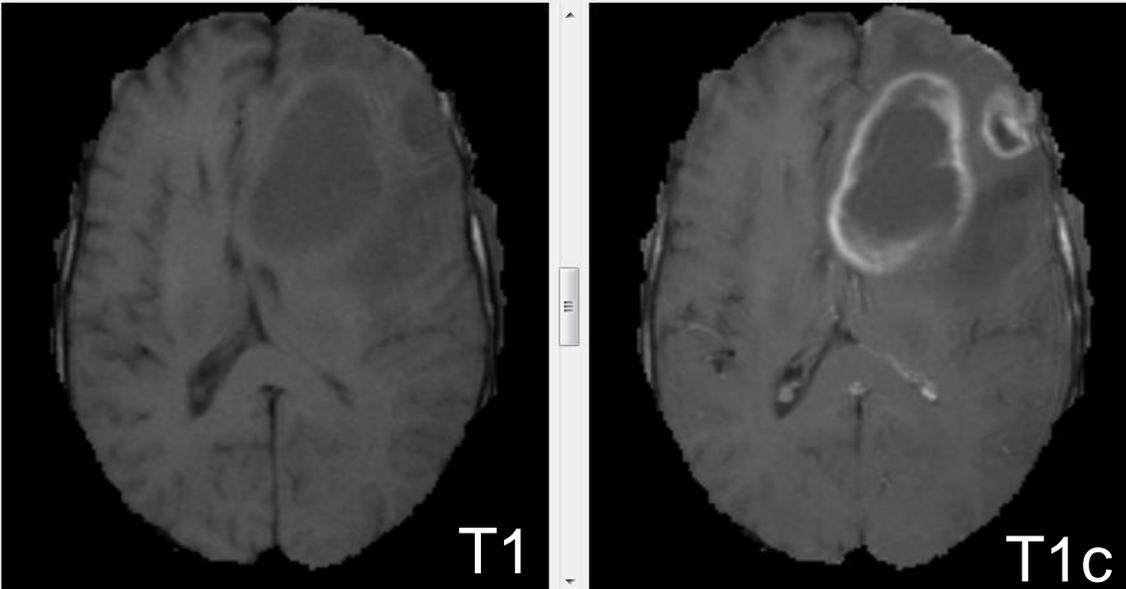
*家族性AD発症のリスク遺伝子: APP, PSEN1, PSEN2

例② Glioblastoma multiforme (GBM)

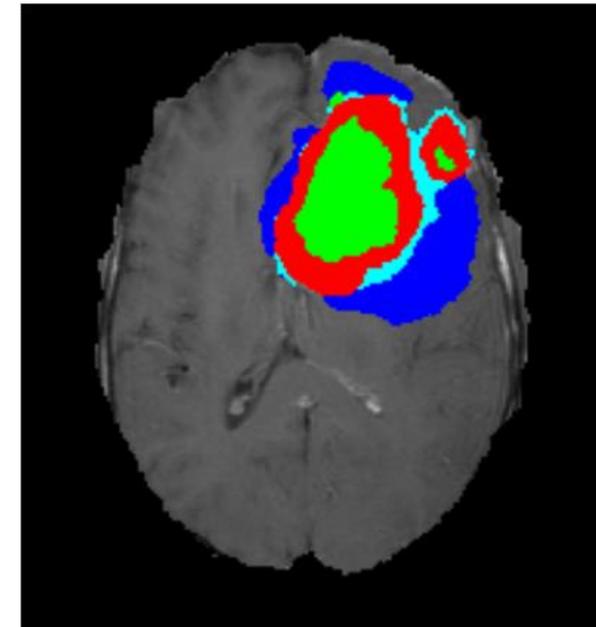
- グリオブラストーマ, 膠芽腫(こうがしゅ)
- WHO グレード4の神経膠腫(グリオーマ)
- 平均生存時間は1年 (Krex *et al* 2007).
- 腫瘍の部位により手術摘出度があがる事などもあり, 部位が生存期間に影響しうる.
- 膠芽腫において, MRIによる画像診断は標準的である.

脳腫瘍における脳画像

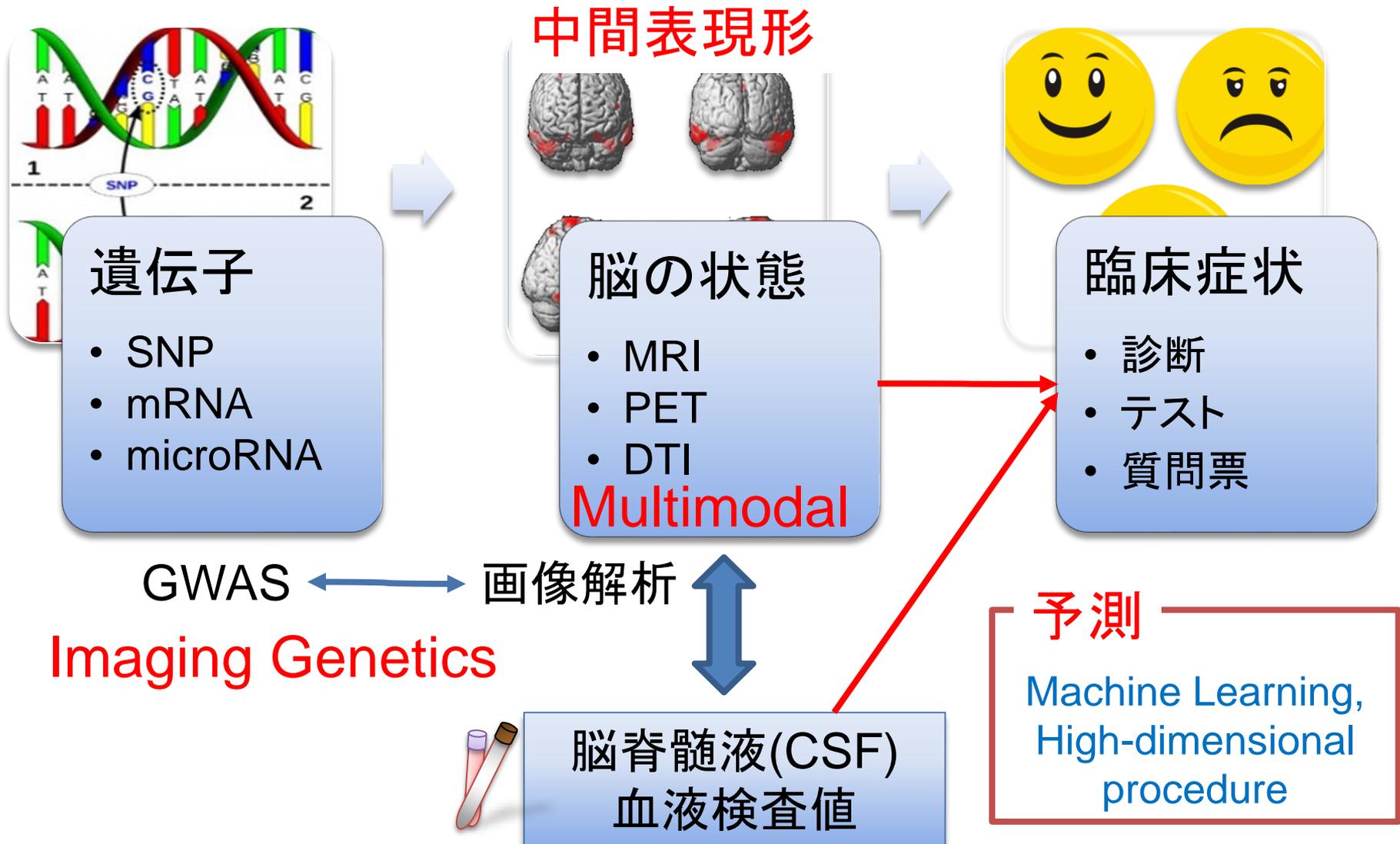
Kawaguchi (2017)



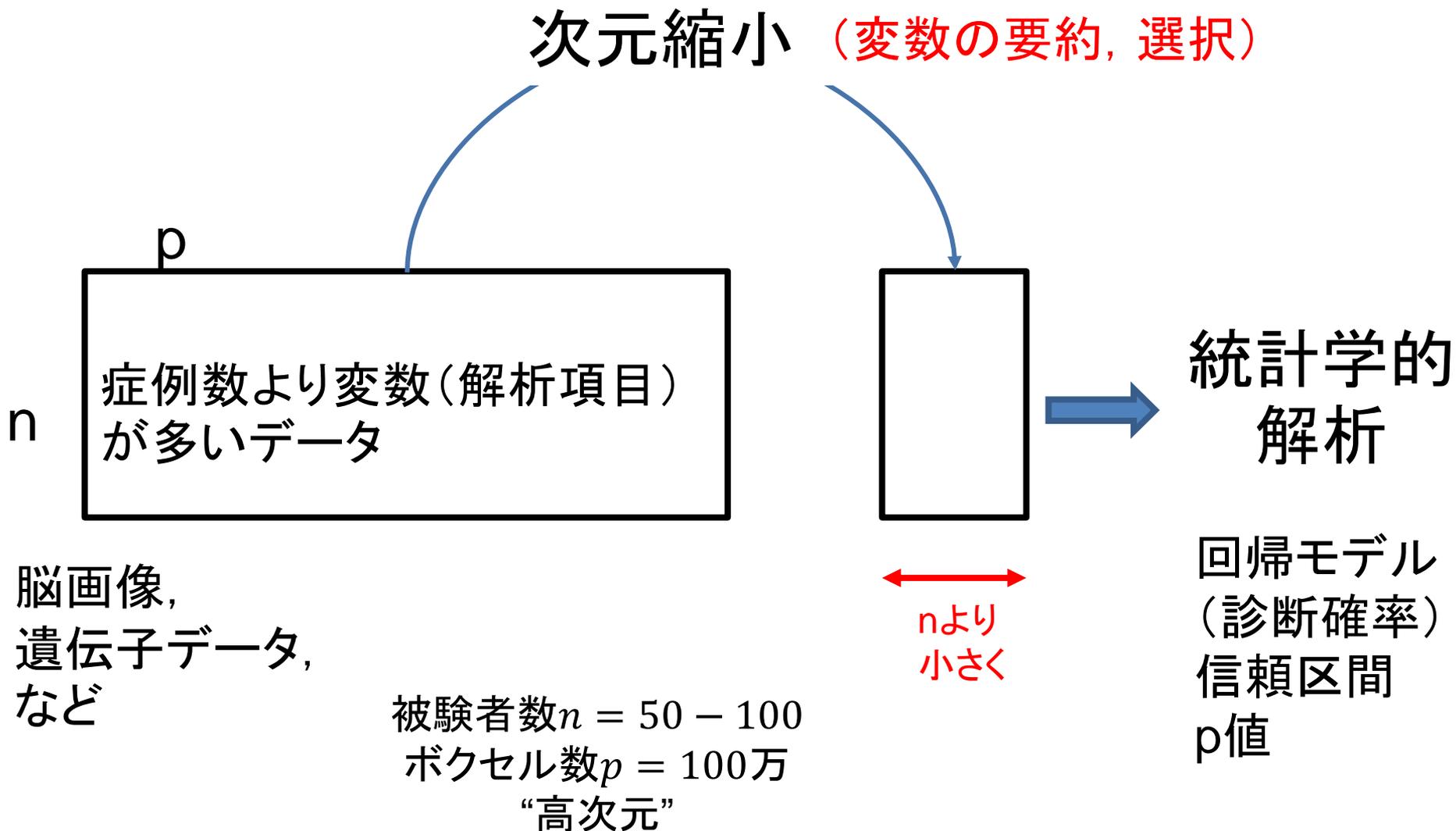
- One axial slice of an MR image of a GBM patient.
 - » 浮腫(dark blue),
 - » 壊死(green),
 - » enhancing active tumor (red)
 - » non-enhancing active tumor (light blue)



MUBIG (Multimodal Brain Imaging Genetics)



次元縮小



Biostatistics

Issues Advance articles Submit ▾ Purchase Alerts About ▾ All Biostatistics



Volume 18, Issue 4
October 2017

Supervised multiblock sparse multivariable analysis with application to multimodal brain imaging genetics

Atsushi Kawaguchi ✉, Fumio Yamashita

Biostatistics, Volume 18, Issue 4, 1 October 2017, Pages 651–665, <https://doi.org/10.1093/biostatistics/kxx011>

Published: 19 March 2017 **Article history** ▾



msma: Multiblock Sparse Multivariable Analysis

Several functions can be used to analyze multiblock multivariable data. If the input is a single matrix, then principal components analysis (PCA) is implemented. If the input is a list of matrices, then multiblock PCA is implemented. If the input is two matrices, for exploratory and objective variables, then partial least squares (PLS) analysis is implemented. If the input is two lists of matrices, for exploratory and objective variables, then multiblock PLS analysis is implemented. Additionally, if an extra outcome variable is specified, then a supervised version of the methods above is implemented. For each method, sparse modeling is also incorporated. Functions for selecting the number of components and regularized parameters are also provided.

Version: 1.0
 Depends: [mvtnorm](#)
 Published: 2018-03-01
 Author: Atsushi Kawaguchi
 Maintainer: Atsushi Kawaguchi <kawa_a24 at yahoo.co.jp>
 License: [GPL-2](#) | [GPL-3](#) [expanded from: GPL (≥ 2)]
 NeedsCompilation: no
 Citation: [msma citation info](#)
 CRAN checks: [msma results](#)

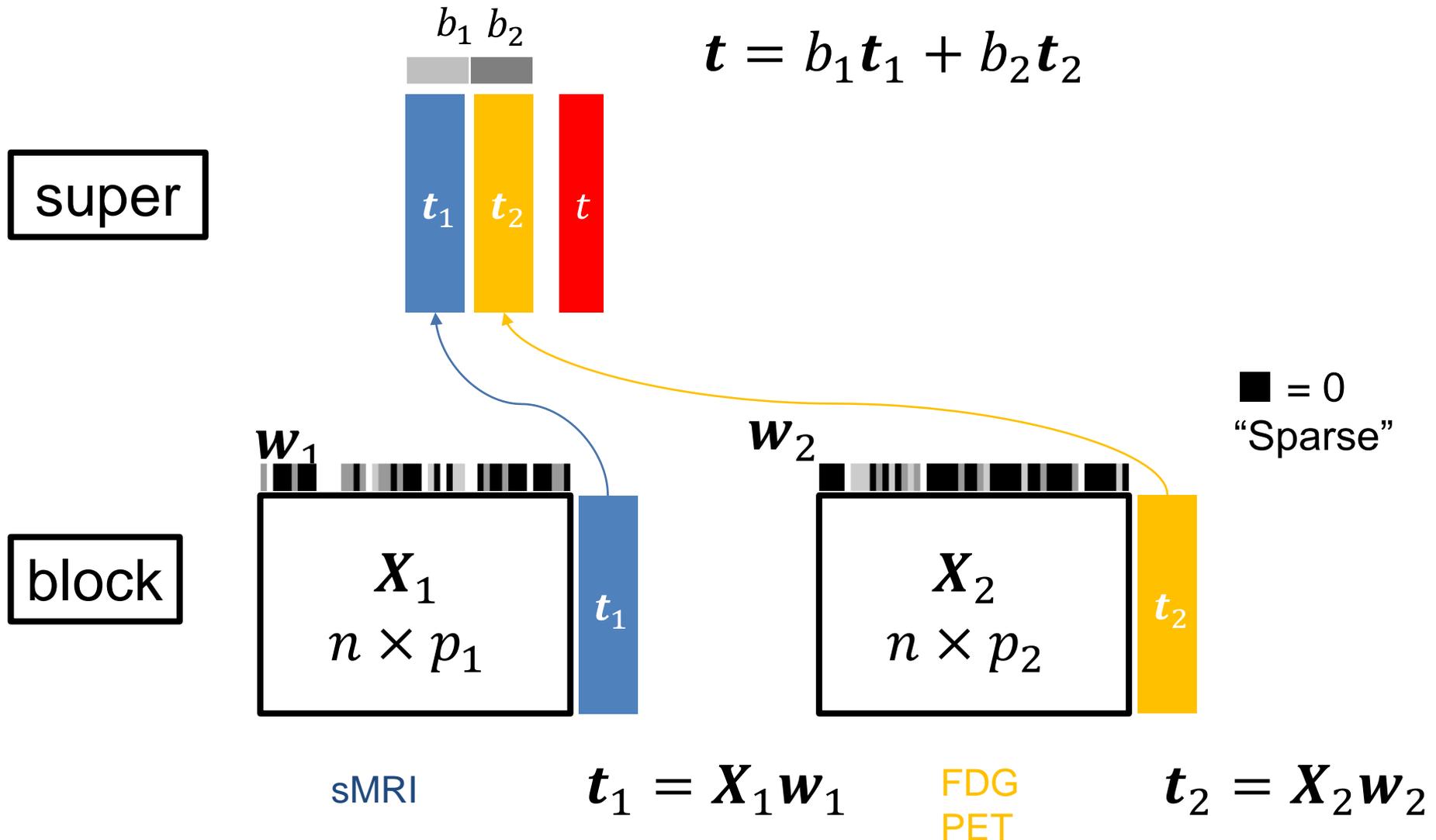
①階層(マルチブロック)

②成分(スコア)

③教師付き

④msma パッケージ(R)

マルチブロック成分法



記法

- n 人の被験者を考える.
- \mathbf{X}_m $n \times p_m$ ($m = 1, 2, \dots, M$) 行列.
 - » M はブロック(モダリティ)数
- \mathbf{X} のスコア表現.

$$\text{スーパースコア} \quad \mathbf{t} = \sum_{m=1}^M b_m \mathbf{X}_m \mathbf{w}_m = \sum_{m=1}^M b_m \mathbf{t}_m \quad \text{ブロックスコア}$$

- 重みベクトル

» $\mathbf{w} = (\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_M)^\top$ ブロック重みベクトル

» $\mathbf{b} = (b_1, b_2, \dots, b_M)^\top$ スーパー重みベクトル

スコアの分散教師付き最大化

$$\mathbf{t} = \sum_{m=1}^M b_m \mathbf{X}_m \mathbf{w}_m, \quad \mathbf{Z} : n\text{次元アウトカムベクトル}$$

- 次の関数を最大化

$$L(\mathbf{b}, \mathbf{w}) = (1 - \mu) \mathbf{t}^\top \mathbf{t} + \mu \mathbf{t}^\top \mathbf{Z} - \sum_{m=1}^M P_{\lambda_m}(\mathbf{w}_m)$$

- subject to $\|\mathbf{w}_m\|_2^2 = 1$ and $\|\mathbf{b}\|_2^2 = 1$ with $\|\cdot\|_2$ as the L2 norm,
- $0 \leq \mu \leq 1$ is the proportion of the supervision,
- $P_\lambda(x) = 2\lambda|x|$
- $\lambda > 0$ is the regularized parameter that is used to control the sparsity.

アルゴリズム

ブロック/スーパー重み

- Initialize \mathbf{t} with $\|\mathbf{t}\|_2 = 1$.
- Repeat until convergence:
 - » Set $\tilde{\mathbf{w}}_m = h_{\lambda_m}(b_m \mathbf{X}_m^\top \{(1 - \mu)\mathbf{t} + \mu \mathbf{Z}\})$,
where $h_\lambda(y) = \text{sign}(y) (|y| > \lambda)_+$, and
normalize as $\hat{\mathbf{w}}_m = \tilde{\mathbf{w}}_m / \|\tilde{\mathbf{w}}_m\|_2$ ($m = 1, 2, \dots, M$).
 - » Set $\mathbf{t}_m = \mathbf{X}_m \hat{\mathbf{w}}_m$ and $\tilde{b}_m = \mathbf{t}_m^\top \{(1 - \mu)\mathbf{t} + \mu \mathbf{Z}\}$;
then set $\tilde{\mathbf{b}} = (\tilde{b}_1, \tilde{b}_2, \dots, \tilde{b}_M)^\top$ and normalize as $\hat{\mathbf{b}} = \tilde{\mathbf{b}} / \|\tilde{\mathbf{b}}\|_2$.
 - » Set $\mathbf{t} = \sum_{m=1}^M \hat{b}_m \mathbf{X}_m \hat{\mathbf{w}}_m$.
- (Deflation step)
Set $\mathbf{p}_m = \mathbf{X}_m^\top \mathbf{t}_m / \mathbf{t}_m^\top \mathbf{t}_m$ and $\hat{\mathbf{X}}_m = \mathbf{t}_m \mathbf{p}_m^\top$, and $\mathbf{X}_m \leftarrow \mathbf{X}_m - \hat{\mathbf{X}}_m$.

上記1.2.3を繰り返す、各componentのスコアと重みを得る

適用例

US-ADNIデータ 画像 (structural MRI and FDG PET)

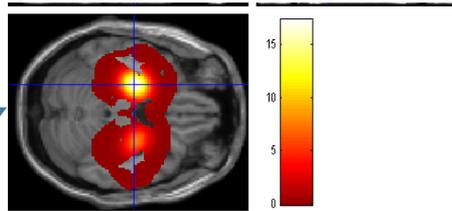
	Dementia	Normal	p
n	52	54	
AGE (mean (sd))	75.41 (7.18)	74.93 (4.89)	0.684
PTGENDER = Male (%)	31 (59.6)	36 (66.7)	0.582
APOE4 (%)			<0.001
0	17 (32.7)	39 (72.2)	
1	29 (55.8)	13 (24.1)	
2	6 (11.5)	2 (3.7)	
PTEDUCAT (mean (sd))	14.19 (3.04)	15.89 (2.99)	0.005
CDRSB (mean (sd))	4.54 (1.73)	0.03 (0.12)	<0.001
ADAS11 (mean (sd))	18.70 (5.63)	6.56 (3.28)	<0.001
ADAS13 (mean (sd))	28.94 (6.30)	10.08 (4.30)	<0.001
MMSE (mean (sd))	23.38 (2.07)	28.87 (1.24)	<0.001

$$Z = 3.17 \times CDR + 0.11 \times ADAS13 - 0.57 \times MMSE$$

where CDR is the clinical dementia rating score,
 ADAS13 is the Alzheimer's disease assessment scale-cognitive subscale, and
 MMSE is the mini-mental state examination score.

二段階次元縮小(脳画像の解析)

結果の表示



復元

■ = 0
"Sparse"

Bw

w

sMRI
 $S_{(n \times N)}$

①基底関数
展開
(構造相関)

縮小画像
 $X_{(n \times p)}$

u

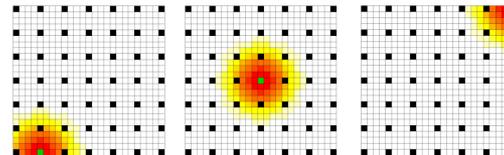
被験者数 $n = 50 - 100$
ボクセル数 $N = 100$ 万
"高次元"

$$X = SB$$

②"スコア化"
(標本相関)

$$u = Xw$$

$$= SBw$$



元データ

Kawaguchi and Yamashita (2017), Kawaguchi (2016, 2017, 2018),
Araki et al. (2013), Yoshida et al. (2013, 2018),

Block weights

Scoreにおいて最も群間差があったcomponent

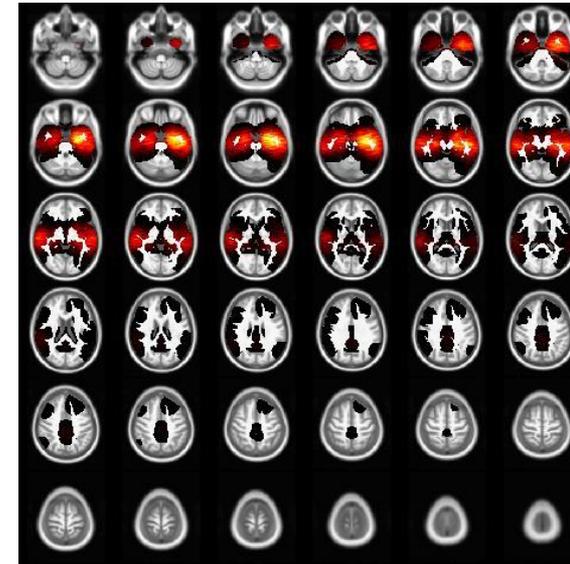
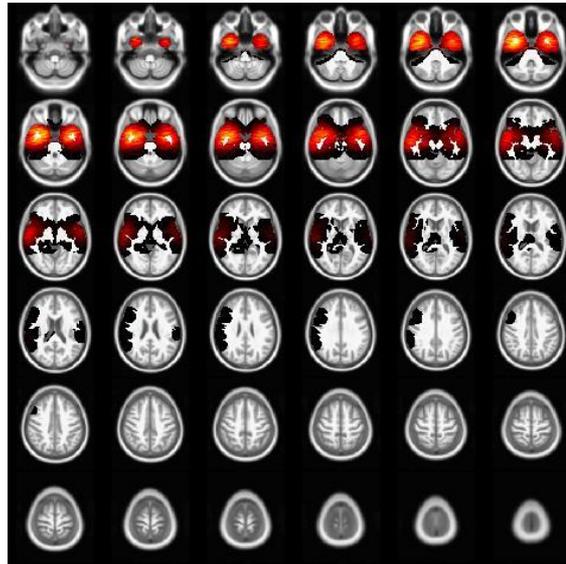
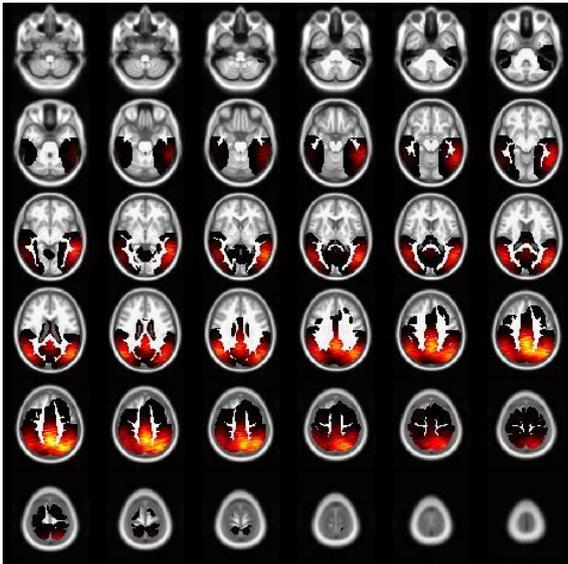
$$(1 - \mu) t^T t + \mu t^T Z$$

$\mu = 0.00$

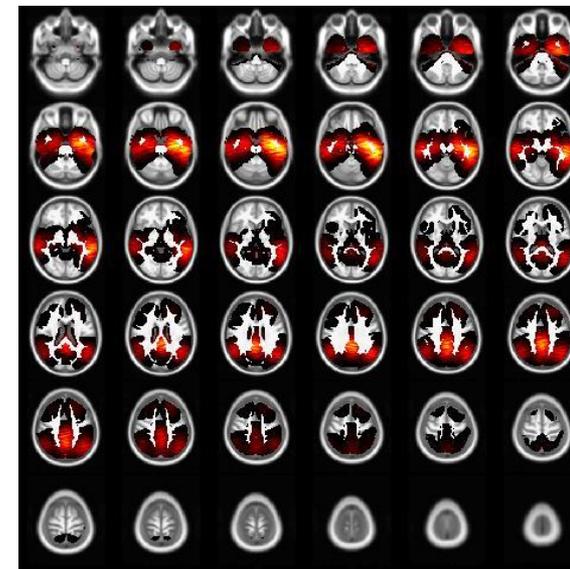
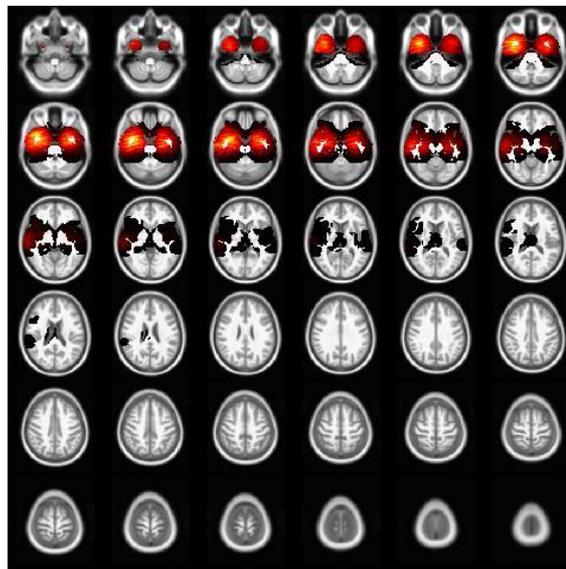
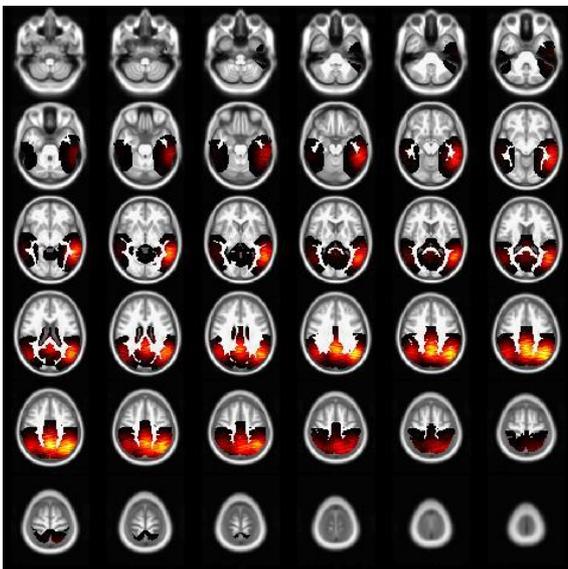
$\mu = 0.50$

$\mu = 1$

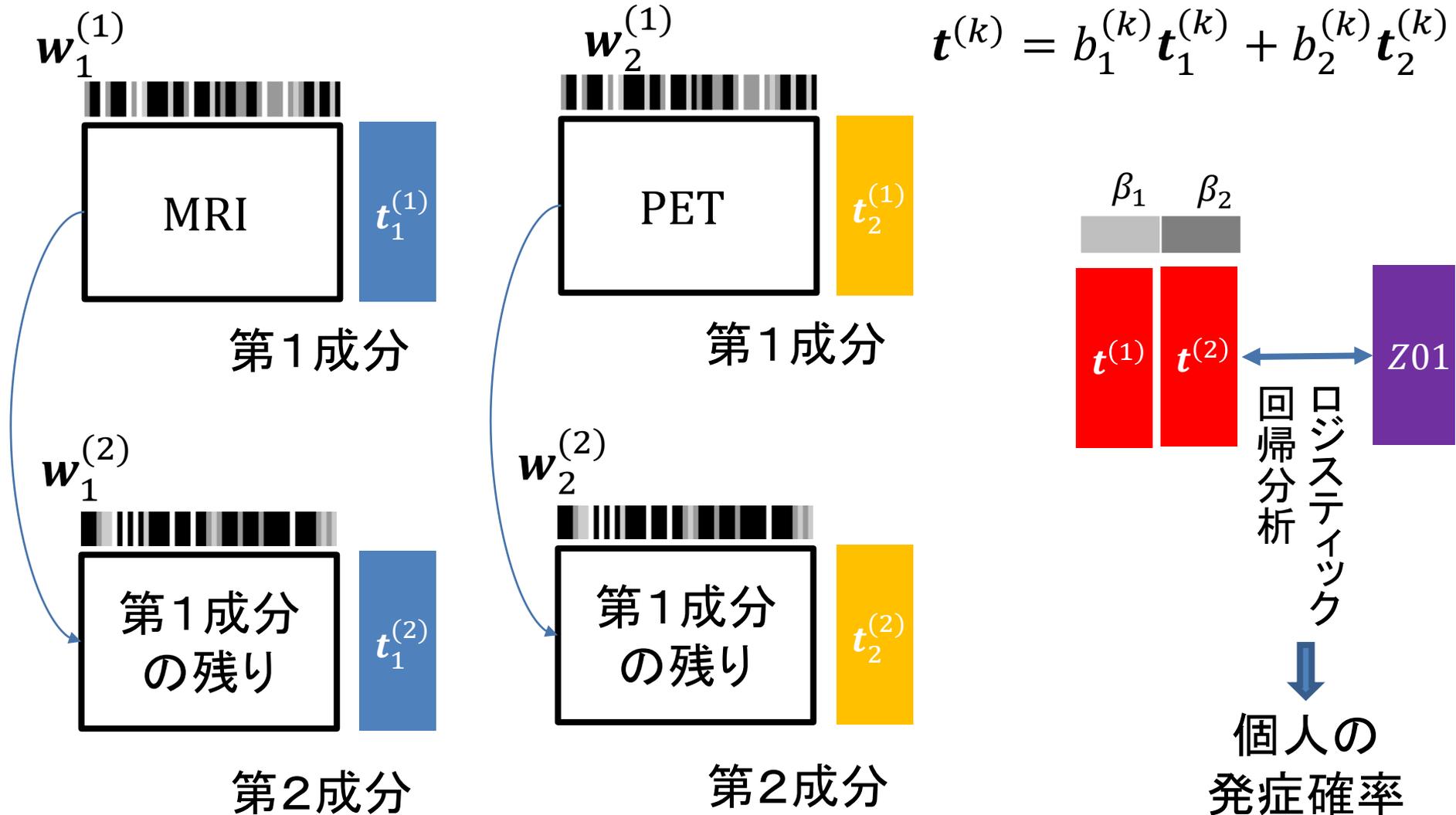
MRI



PET



ロジスティック回帰分析（次元縮小付き）



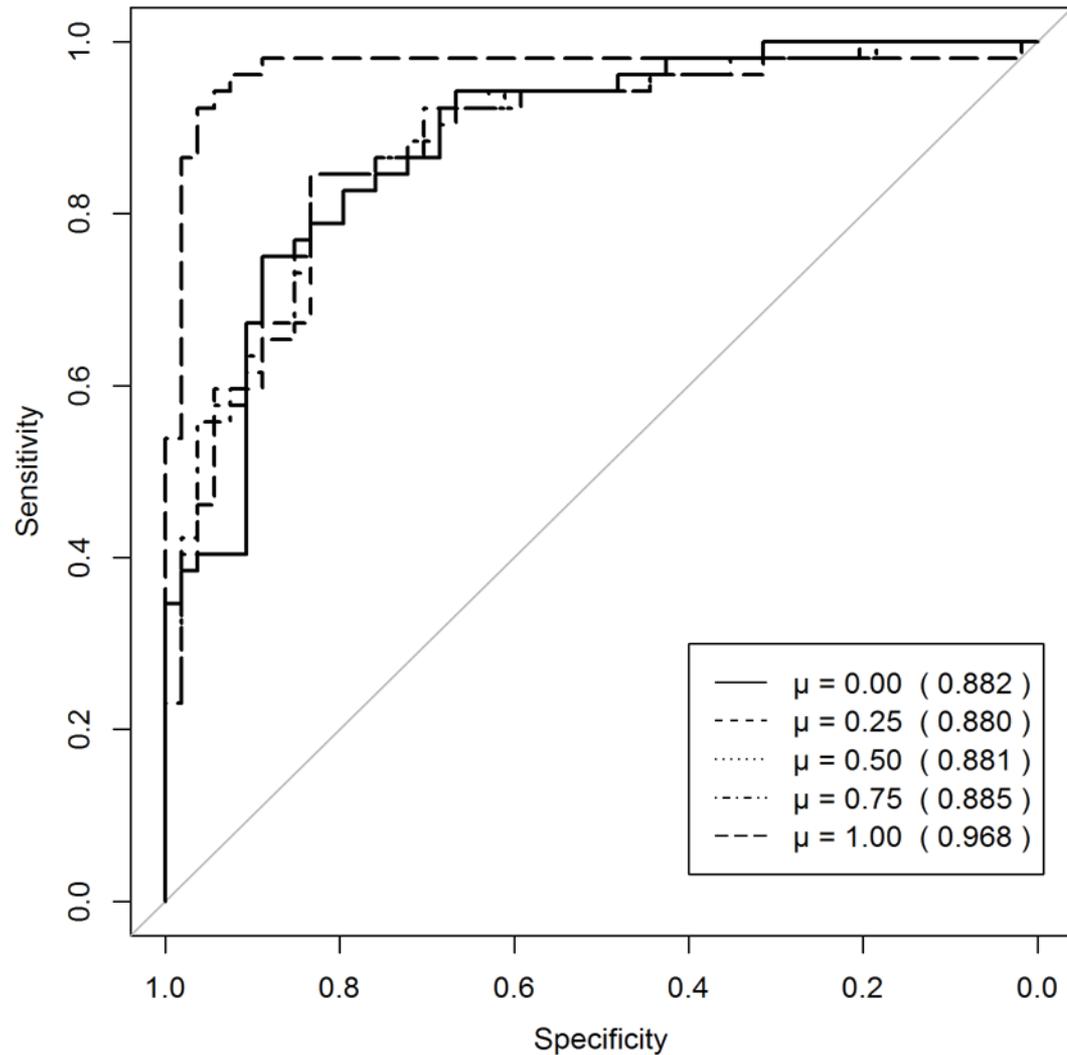
Model fit summary

多重ロジスティック回帰

	$\mu = 0.00$			$\mu = 0.5$			$\mu = 1.0$		
	Estimate	Pr(> z)		Estimate	Pr(> z)		Estimate	Pr(> z)	
comp1	-0.0210	0.0615		0.0832	<0.0001	<0.0001	4.287	0.9982	0.0001
comp2	0.0882	<0.0001	<0.0001	0.0458	0.0031	0.0026	2.555	0.9989	0.0013
comp3	-0.0621	0.0001	0.0001	0.0180	0.0920		4.633	0.9995	0.0027
comp4	-0.0072	0.6126		0.0044	0.7583		1.827	0.9987	
comp5	-0.0424	0.0228	0.0446	0.0430	0.0203	0.0261	3.905	0.9988	
comp6	-0.0364	0.0900	0.0439	0.0403	0.0715	0.0750	4.994	0.9984	
comp7	0.0446	0.0891		0.0510	0.0619	0.0543			
comp8	0.0336	0.2517		-0.0226	0.3816				

ROC

多重ロジスティック回帰 with stepwise selection
10-fold cross-validation



まとめ

➤ マルチブロック成分法

- » ブロックとスーパーの2つの重み/スコア
- » 教師付最適化

➤ 適用例

- » PLSまで拡張 (kawaguchi and Yamashita 2017)
 - > 画像(sMRI)-遺伝子(SNP) 解析
 - > 画像(Amyloid)-画像(sMRI) 解析
- » 同一モダリティの経時測定 (kawaguchi 2018)
- » fMRI(時系列画像)解析 ← 深層ブロックの適用

Reference

- Arbabshirani, M. R., Plis, S., Sui, J., & Calhoun, V. D. (2017). Single subject prediction of brain disorders in neuroimaging: Promises and pitfalls. *NeuroImage*, 145, 137-165.
- Kawaguchi A, Yamashita F (2017). Supervised Multiblock Sparse Multivariable Analysis with Application to Multimodal Brain Imaging Genetics. *Biostatistics*, 18(4) 651-665.
- Kawaguchi A (2017). Supervised Dimension Reduction Methods for Brain Tumor Image Data Analysis. In *Frontiers of Biostatistical Methods and Applications in Clinical Oncology*, Matsui S and Crowley J. (Ed.), Springer, 401--412.
- Kawaguchi A (2018). Supervised sparse components analysis with application to brain imaging data. In *Neuroimaging*, Golubic SJ (Ed.), IntechOpen in press.
- Yoshida H, Kawaguchi A, Yamashita F, Tsuruya K (2018). The utility of a network-based clustering method for dimension reduction of imaging and non-imaging biomarkers predictive of Alzheimer's disease. *Scientific Reports*, 8, 2807.