

プリオン病サーベイランスデータの管理・運用の研究

研究分担者：東海大学医学部臨床薬理学

金谷泰宏

表1 機械学習モデルによる病型別の診断確率

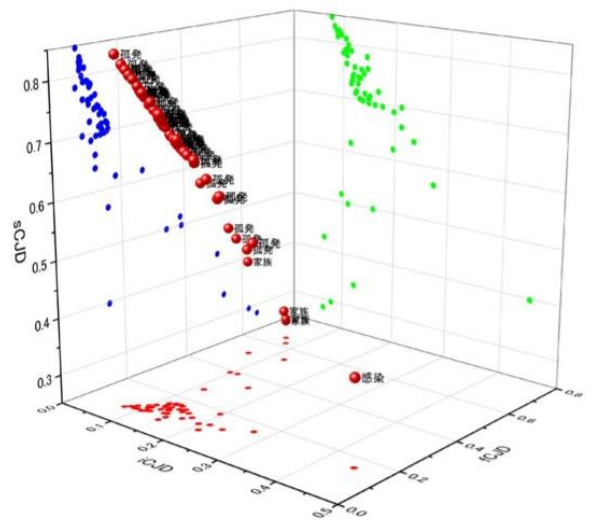
Prediction performance (AUC)				
Fold	1: 孤発	2: 遺伝性	3: 感染性 (乾燥硬膜移植)	Average
0	0.868	0.951	0.723	0.847
1	0.885	0.957	0.812	0.885
2	0.850	0.911	0.707	0.823
3	0.865	0.932	0.692	0.830
4	0.893	0.992	0.719	0.868
AUC, Mean	0.872	0.949	0.731	0.851
AUC, SD	0.017	0.030	0.047	0.026

表2 病型別における診断への項目の影響度

「1はい 2いいい」の場合、「はい」であればweightはマイナス値となる

病型	score	weight	Feature	score	weight	Feature	score	weight
孤発型	0.5374	-0.1491	プリオン蛋白遺伝子検査実施	0.8910	-0.2417	プリオン蛋白遺伝子検査実施	0.7220	-0.2176
孤発型	0.5254	-0.0227	流行性痴呆又は変性障害	0.8370	-0.2903	diffusion 又は FLAIRで高信号	0.8370	-0.3095
孤発型	0.3855	-0.1879	NSE増減	0.2230	-0.1708	NSE増減	0.2735	-0.2489
孤発型	0.2490	-0.1533	ミオクローヌス	0.1879	-0.0227	コドシ219の多型 3.0	0.2294	-0.6054
孤発型	0.2149	0.4524	プリオン蛋白遺伝子変異有無	0.1879	-0.0227	コドシ219の多型 3.0	0.1931	-0.0478
孤発型	0.1901	0.5945	プリオン蛋白遺伝子変異有無	0.1859	0.1170	プリオン蛋白遺伝子変異有無	0.1711	0.4932
孤発型	0.1440	-0.1878	コドシ219の多型 2.0	0.1325	0.2193	細胞数増減	0.1095	0.0824
孤発型	0.0967	0.0790	精神症状	0.1001	0.0695	視覚異常	0.0750	0.0250
孤発型	0.0757	-0.0753	小脳症状ふらつき	0.0736	0.0389	異常運動の増進化	0.0747	-0.0789
孤発型	0.0749	-0.2335	無動・舞臺状態	0.0685	-0.0200	PSD	0.0492	-0.0482
孤発型	0.0688	-0.0908	diffusion 又は FLAIRで高信号	0.0604	0.0809	性別	0.0427	-0.0016
孤発型	0.0666	0.0908	硬膜移植	0.0530	0.0883	申請年度 西暦	0.0505	-0.0363
孤発型	0.0505	-0.0309	性別	0.0476	-0.0159	硬膜移植	0.0471	-0.1269
孤発型	0.0504	-0.0690	申請年度 西暦	0.0440	-0.0386	精神症状	0.0457	-0.0023
孤発型	0.0501	-0.0300	細胞数	0.0397	-0.0879	コドシ219の多型 3.0	0.0364	-0.0250
孤発型	0.0480	0.0722	難病外診断	0.0334	0.0240	蛋白質増減	0.0361	0.0238
孤発型	0.0365	-0.1068	コドシ219の多型 3.0	0.0239	-0.1885	コドシ219の多型 1.0	0.0226	-0.2249
孤発型	0.0196	0.0168	生年月日	0.0221	-0.0190	難病外診断	0.0165	0.0070
孤発型	0.0186	0.0346	蛋白質増減	0.0205	0.0424	生年月日	0.0155	0.0048
孤発型	0.0117	-0.0088	視覚異常	0.0179	0.0244	細胞数	0.0152	0.0539
孤発型	0.0079	0.0344	発症時年齢	0.0131	-0.0182	流行性痴呆又は変性障害	0.0148	0.0748
孤発型	0.0071	0.0095	異常運動の増進化	0.0097	0.0204	発症時年齢	0.0112	-0.0405
孤発型	0.0067	0.0328	コドシ219の多型 3.0	0.0073	-0.0222	使用硬膜製品名	0.0074	-0.1576
孤発型	0.0052	-0.1882	蛋白質	0.0040	-0.0792	コドシ219の多型 2.0	0.0048	0.0048
孤発型	0.0043	0.0110	難病外診断	0.0056	-0.0047	蛋白質	0.0048	-0.0776
孤発型	0.0039	-0.0184	細胞数増減	0.0051	0.0487	コドシ219の多型 1.0	0.0041	-0.0813
孤発型	0.0030	0.0248	CT MRIで脳萎縮	0.0039	-0.0261	CT MRIで脳萎縮	0.0033	-0.0069
孤発型	0.0025	-0.0522	コドシ219の多型 3.0	0.0035	-0.1240	小脳症状ふらつき	0.0030	0.0059
孤発型	0.0014	-0.0112	発症時年齢	0.0023	-0.0084	コドシ219の多型 3.0	0.0012	-0.0054
孤発型	0.0005	-0.0383	使用硬膜製品名	0.0016	-0.2418	発症時年齢	0.0010	0.0180
孤発型	0.0003	-0.0021	コドシ219の多型 1.0	0.0012	0.0578	NSE量	0.0004	-0.0015
孤発型	0.0001	-0.0361	NSE量	0.0009	0.0087	難病外診断	0.0002	0.0083
孤発型	0.0000	-0.0088	コドシ219の多型 1.0	0.0008	-0.0214	コドシ219の多型 3.0	0.0001	-0.0049

図1 機械学習モデルによる三次元マッピング



解説

- 個人票情報を機械学習させることで診断予測モデルを作成した(表1)。診断予測モデルの精度は、全体でAUC 0.851と極めて高い値を示した。なお、感染性CJDではAUC=0.731と予測精度が低い傾向が示された。
- 機械学習として、Point wise linearモデルを用いることで、病型別に診断に大きく影響する因子を抽出することができた(表2) これにより、3病型の分類に際して、いかなる要素が大きく関与しているかを明らかにした。3病型に共通する要素としてプリオン蛋白遺伝子検査実施が認められた。一方で、ミオクローヌスは、孤発型、遺伝型で上位に位置付けられたが(0.5374, 0.1879)、感染型では低い位置付けとなった(0.0364)。なお、視覚異常は感染性で0.075と高く、遺伝性で0.0179、孤発性で<0.0001と極めて低い傾向が示された。
- 53例をランダムに選択し、機械学習モデルにより診断確率を算出した後、三次元データプロットを行った(図1)。これにより、sCJD、fCJD、iCJDの病型間の関係性を明らかにすることができた。