



## 演題名：新規スクリーニング糖尿病患者の受診勧奨後未受診の予測モデル構築

演者名：岡田 啓<sup>1)</sup>、橋本 洋平<sup>2,3)</sup>、後藤 匡啓<sup>2,4)</sup>、山口 聡子<sup>1)</sup>、大野 幸子<sup>5)</sup>、倉川 佳世<sup>1)</sup>、大塚 雄介<sup>1)</sup>、平岡 信歩<sup>1)</sup>、百瀬 彬<sup>1)</sup>、南学 正臣<sup>6)</sup>、山内 敏正<sup>7)</sup>、康永 秀生<sup>2)</sup>、門脇 孝<sup>1,7,8)</sup>  
<sup>1)</sup>東京大学大学院医学系研究科糖尿病・生活習慣病予防講座、<sup>2)</sup>臨床疫学・経済学、<sup>3)</sup>眼科、<sup>4)</sup>TXPメディカル、<sup>5)</sup>イートロス医学講座、<sup>6)</sup>腎臓・内分泌内科、<sup>7)</sup>糖尿病・代謝内科、<sup>8)</sup>虎の門病院

### Background

- ◆ 糖尿病の治療へのアドヒアランスの低さは、入院率や医療コストの増加を招くことが知られており、受診して加療を受けることは重要である。
- ◆ しなしながら、糖尿病スクリーニング後の受診率は低い。実際、過去の報告によると、糖代謝異常で受診勧奨を受けた患者の内、35%しか受診していなかった。
- ◆ 実際、受診中断の因子は数々の研究があるが、受診勧奨後未受診の因子などを検討した論文は殆ど無い。
- ◆ 糖尿病の診断基準が2010年に変わり、HbA1cが組み込まれたという変化、機械学習などの新たな手法の出現の変化もあり、
- ◆ より「良く」「効率的に」糖尿病受診勧奨後未受診を予測すれば政策立案に役立つ可能性がある。
- ◆ 今回我々は、JMDCデータベースと機械学習を用いて、受診勧奨後の未受診予測モデルの構築を試みた。

### Methods

- ◆ データソース：JMDCデータベース
- ◆ 組み入れ基準：新規スクリーニング糖尿病患者（定義：過去1年に糖尿病診療を受けておらず、HbA1c 6.5%以上かつ空腹時血糖値 126 mg/dL以上を満たす新規糖尿病患者）
- ◆ 除外基準：受診勧奨後6ヶ月以内に保険離脱した例、検査値（BMI、腎機能、尿検査、ウエスト周囲長、脂質検査など）が欠測を持つ例
- ◆ 未受診の定義：健診から6ヶ月以上糖尿病関連レセプトが発生していない 追跡期間：健診から6ヶ月
- ◆ 統計解析
  - ・ 全体をまず、4:1に分けてモデル構築コホートとモデル検証コホートに分割
  - ・ モデル構築コホートでは、候補変数の中から、Lasso回帰の1SE modelで変数選択をして未受診の予測モデル構築
  - 変数の候補としては、性別、年齢、血液学的マーカー（糖尿病、脂質異常症、肝機能異常など）、特定健診でのアンケート項目、過去12ヶ月の医療機関受診頻度（月/年）、他の薬剤処方歴など
  - ・ モデル検証コホートにて、機械学習モデルと既存モデル（Diabetes Res Clin Pract 2014;105:176-184）のc統計量をDeLongテストで検定して予測能を比較

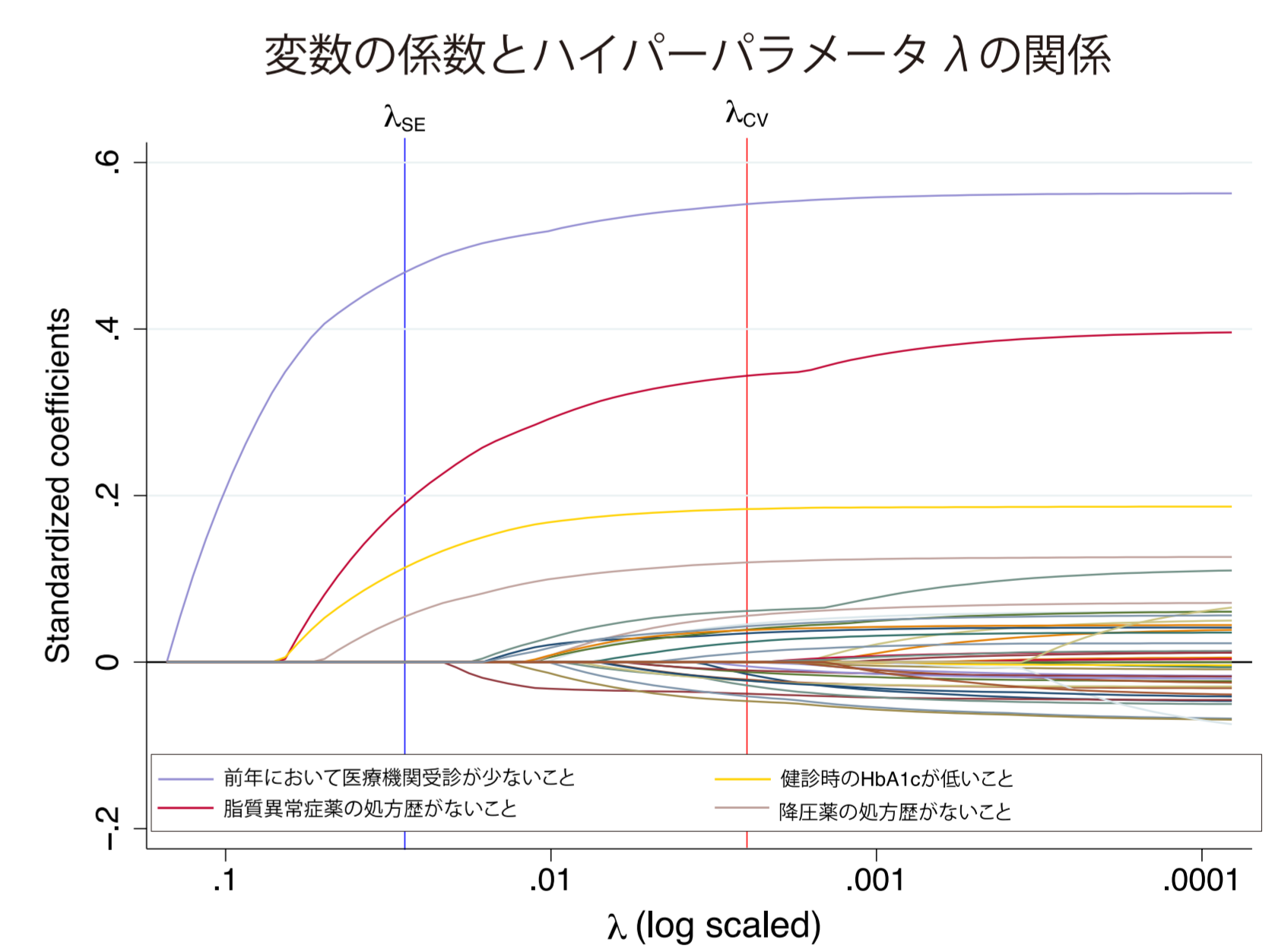
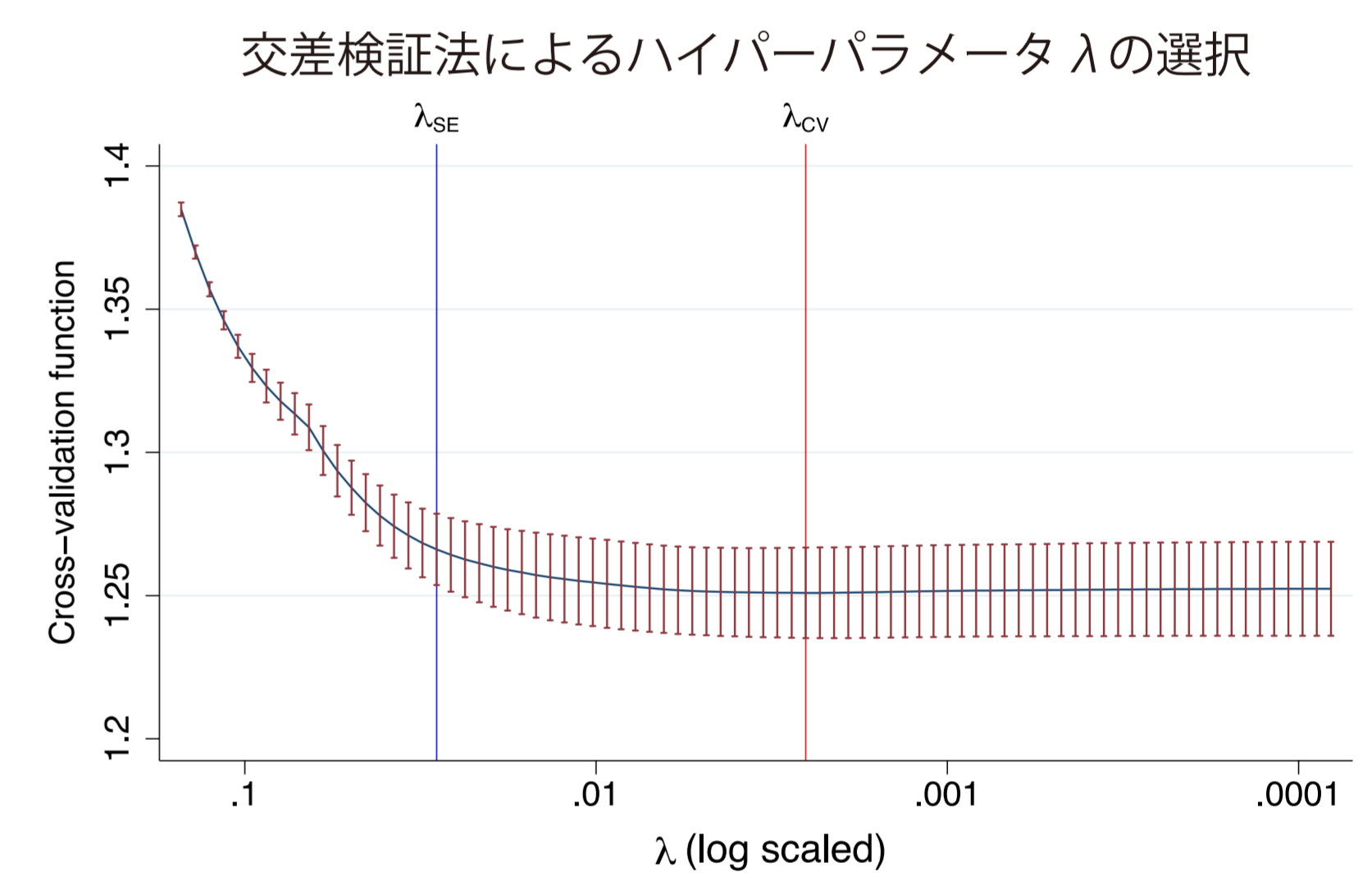
### Results

- ◆ 11,023名が組み入れ基準を満たした。除外が378名で、解析対象者は10,645名
- ◆ 背景としては、未受診群には受診群と比較して、男性が多く、やや若年であった。また、喫煙者が多く、心血管病の既往者が少なかった。さらに、尿蛋白定性陽性が少なく、空腹時血糖値やHbA1c値がやや低めだった。糖尿病以外の生活習慣病関連の処方率も未受診群が低かった。

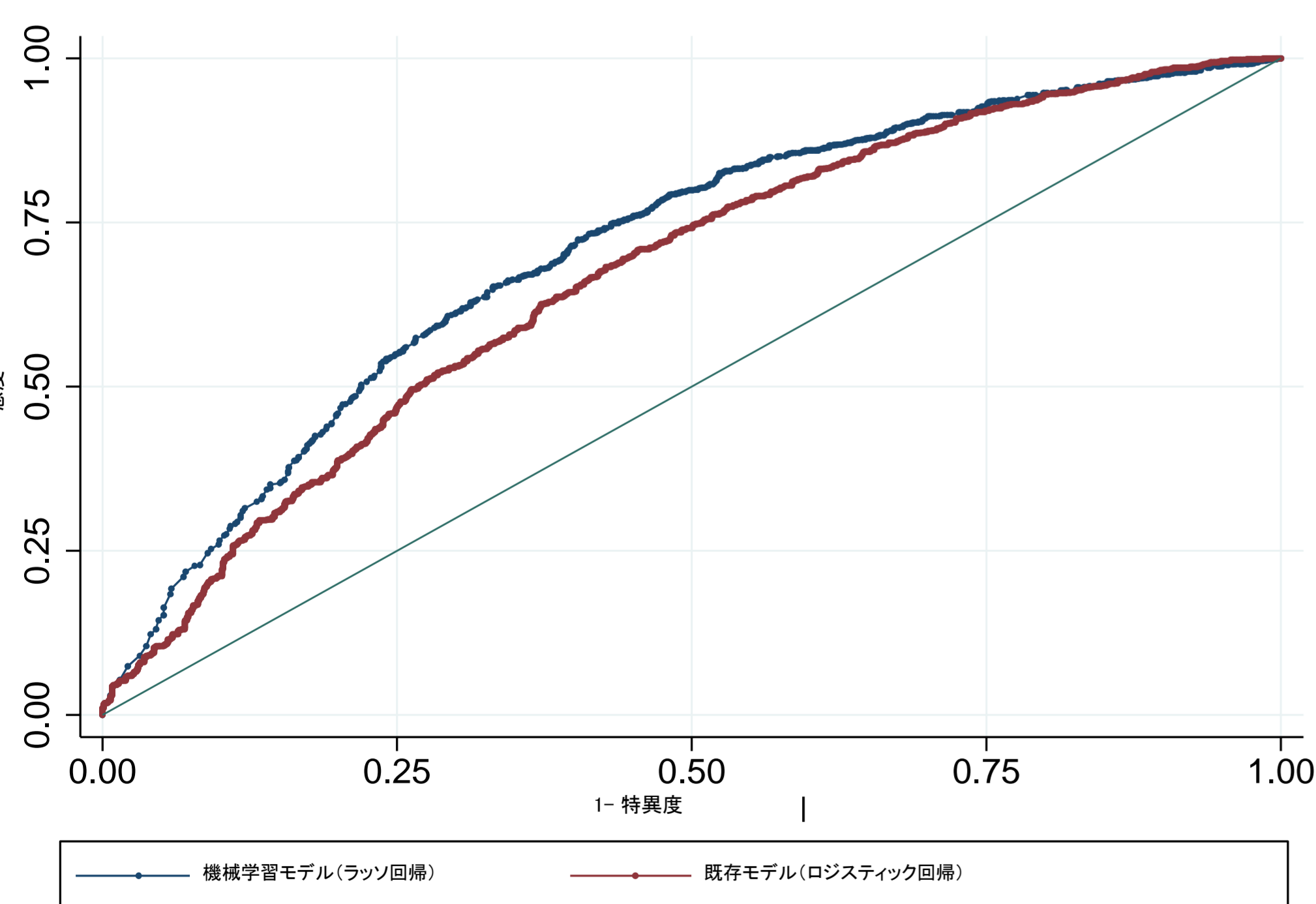
表：健診後に受診した群と受診していない群の背景情報

変数	カテゴリ	受診群 N=5,195	未受診群 N=5,450	P値
性別	男性	4125 (79.4%)	4452 (81.7%)	<0.001
年齢, 中央値 (IQR)		53.0 (48.0, 58.0)	51.0 (46.0, 57.0)	<0.001
Body mass index (kg/m <sup>2</sup> )	<18.5	51 (1.0%)	35 (0.6%)	0.092
	18.5-24.9	1763 (33.9%)	1851 (34.0%)	
	25.0-29.9	2282 (43.9%)	2310 (42.4%)	
	≥30.0	1099 (21.2%)	1254 (23.0%)	
ウエスト周囲長	男性: <85 cm, 女性: <90 cm	1460 (28.1%)	1535 (28.2%)	0.94
	男性: ≥85 cm, 女性: ≥90 cm	3735 (71.9%)	3915 (71.8%)	
血圧	正常	3258 (62.7%)	3338 (61.2%)	0.12
	1度高血圧	1354 (26.1%)	1420 (26.1%)	
	2度高血圧	435 (8.4%)	512 (9.4%)	
	3度高血圧	148 (2.8%)	180 (3.3%)	
喫煙歴	非喫煙者	3433 (66.1%)	3218 (59.0%)	<0.001
	喫煙者	1762 (33.9%)	2232 (41.0%)	
飲酒頻度	めったに/全く飲まない	2179 (41.9%)	2103 (38.6%)	0.002
	時々/習慣的に	1639 (31.5%)	1798 (33.0%)	
心血管病既往		224 (4.3%)	112 (2.1%)	<0.001
保健指導の希望の有無	+	1767 (34.0%)	1716 (31.5%)	0.005
行動変容ステージ	無関心期	812 (15.6%)	959 (17.6%)	<0.001
	関心期	1986 (38.2%)	2166 (39.7%)	
	準備期	941 (18.1%)	910 (16.7%)	
	実行期	545 (10.5%)	582 (10.7%)	
十分な睡眠が取れているか	-	3023 (58.2%)	3161 (58.0%)	0.84
	+	2172 (41.8%)	2289 (42.0%)	

変数	カテゴリ	受診群 N=5,195	未受診群 N=5,450	P値
尿蛋白陽性		672 (12.9%)	598 (11.0%)	0.002
尿糖陽性		1630 (31.4%)	1439 (26.4%)	<0.001
空腹時血糖値 (mg/dL), 中央値 (IQR)		145.0 (134.0, 176.0)	141.0 (132.0, 164.0)	<0.001
HbA1c (%), 中央値 (IQR)		7.3 (6.8, 8.5)	7.1 (6.7, 8.0)	<0.001
中性脂肪 (mg/dL)	<150	2675 (51.5%)	2745 (50.4%)	0.51
	150-299	1862 (35.8%)	1998 (36.7%)	
	≥300	658 (12.7%)	707 (13.0%)	
LDL-コレステロール (mg/dL)	<120	1620 (31.2%)	1449 (26.6%)	<0.001
	120-139	1224 (23.6%)	1207 (22.1%)	
HDL-コレステロール (mg/dL)	≥140	2351 (45.3%)	2794 (51.3%)	0.27
	<40	741 (14.3%)	819 (15.0%)	
ヘモグロビン (g/dL), 中央値 (IQR)		15.4 (14.6, 16.2)	15.5 (14.7, 16.3)	<0.001
血清尿酸値 (mg/dL), 中央値 (IQR)		5.6 (4.7, 6.6)	5.8 (4.9, 6.7)	<0.001
推定糸球体濾過量 (ml/min/1.73m <sup>2</sup> ), 中央値 (IQR)		78.7 (68.9, 90.0)	79.4 (70.1, 90.7)	<0.001
メタボリックシンドローム		3697 (71.2%)	3868 (71.0%)	0.83
Fatty liver index, 中央値 (IQR)		91.7 (10.3, 99.9)	93.3 (13.3, 99.9)	0.15
保険加入	本人	4060 (78.2%)	4460 (81.8%)	<0.001
	被扶養者	1135 (21.8%)	990 (18.2%)	
降圧薬処方		1578 (30.4%)	667 (12.2%)	<0.001
脂質異常症薬処方		1018 (19.6%)	335 (6.1%)	<0.001
高尿酸血症薬処方		417 (8.0%)	178 (3.3%)	<0.001
抗うつ剤処方		188 (3.6%)	107 (2.0%)	<0.001
過去12ヶ月の受診頻度 (月数/年), 中央値 (IQR)		5.0 (1.0, 10.0)	2.0 (0.0, 5.0)	<0.001
前回健診時の受診勧奨の有無		4526 (87.1%)	4737 (86.9%)	0.75

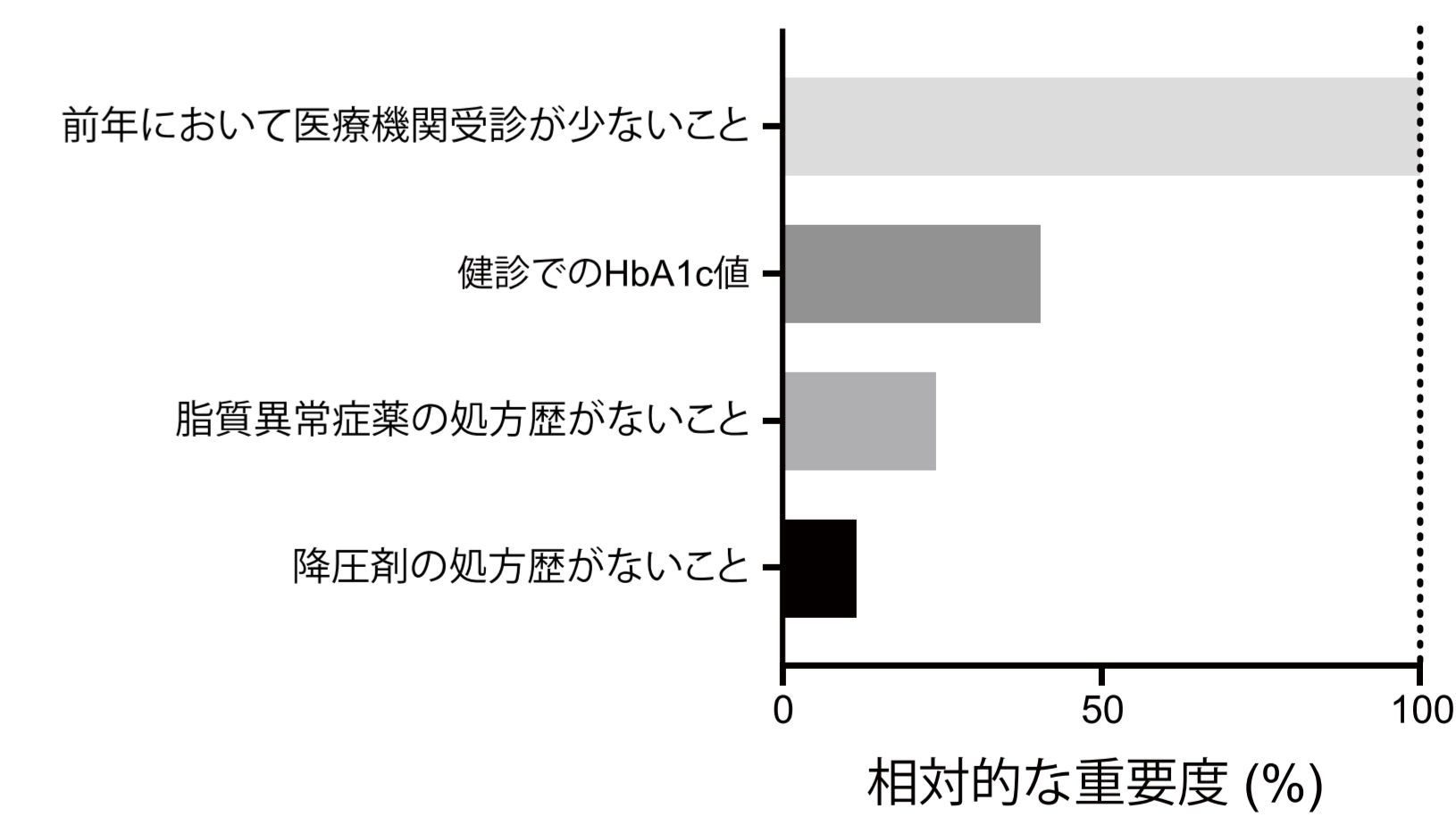


図：既存モデルと機械学習モデルの予測能の比較

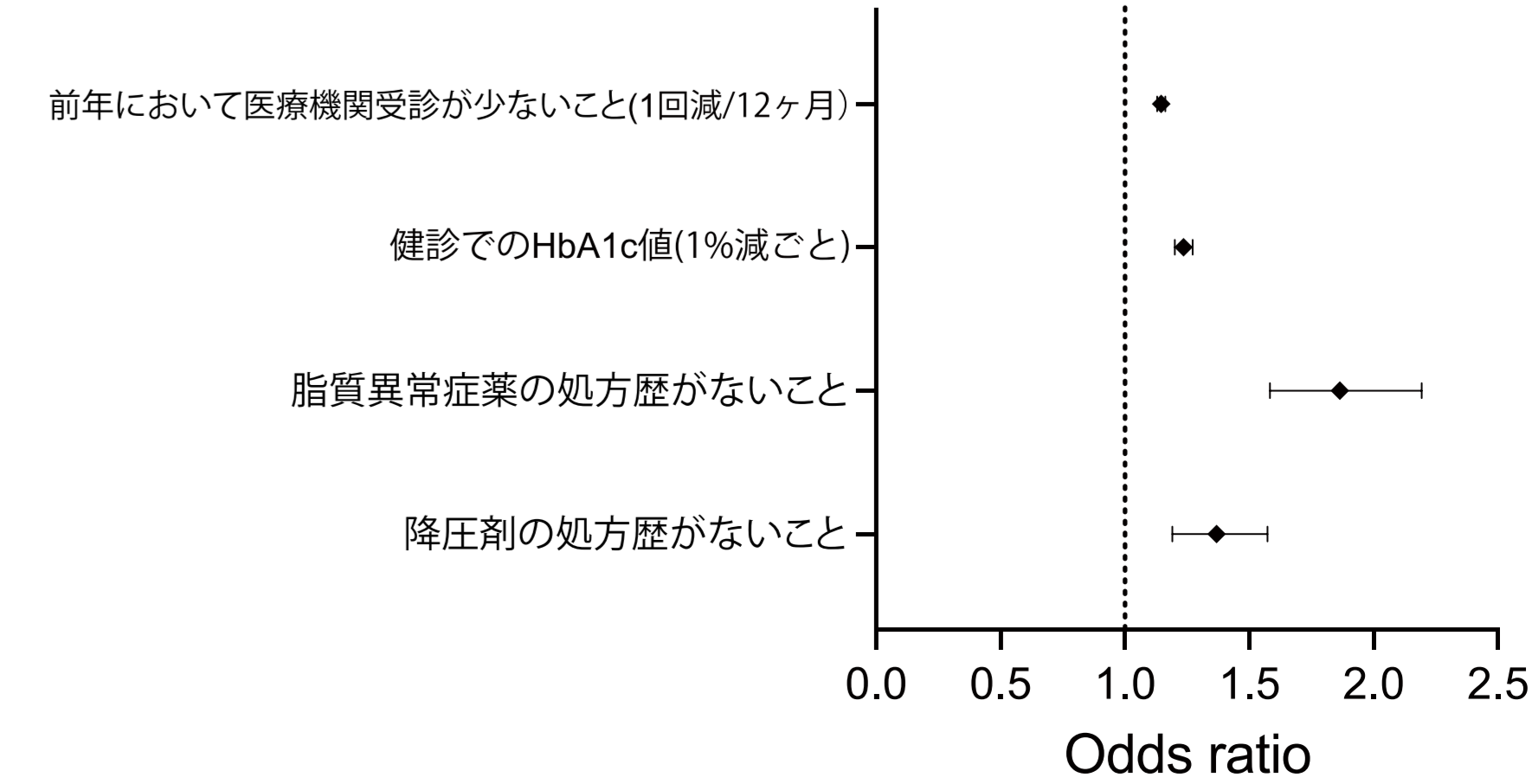


- ◆ モデル構築コホートで、上記合計39因子から交差検証法を行い機械学習モデル（Lasso回帰1標準誤差モデル）を用いて、変数を4つに絞った（右図）。
- ◆ 機械学習モデルによって選ばれた因子は、(1) 過去12ヶ月の受診頻度 (2) HbA1c (3) 脂質異常症薬処方 (4) 降圧薬処方だった。
- ◆ モデル検証コホートを用いた予測能の比較では、既存モデルがc統計量0.67 (95%信頼区間0.65-0.69) 機械学習モデルがc統計量0.71 (95%信頼区間0.69-0.73) であり、DeLong検定にてP値<0.001と、機械学習モデルが予測能が高かった。
- ◆ 本結果は、HbA1c値に関わらず一貫していた（data not shown）。

Lasso回帰モデルにおける変数の重要度



Lasso回帰モデルにおける変数のオッズ比



### Conclusions

- 本研究により、健診での新規糖尿病指摘後に未受診に至る予測を、4因子のみで既存研究の因子より予測能を高めることが出来た。
- 重要な因子は、重要な順に(1) 過去12ヶ月の受診頻度 (2) HbA1c (3) 脂質異常症薬処方 (4) 降圧薬処方だった。
- 医療機関の利用頻度が少ない人はより積極的に受診を促す必要がある可能性がある。

本発表は、本発表抄録登録後、Diabetes Careに受理された内容を説明しています。