20240530　第10回協会けんぽ調査研究フォーラム
@一橋大学　一橋講堂

機械学習を用いた生活習慣病の治療行動予測モデルの構築

研究代表者　丹野高三　岩手医科大学 衛生学公衆衛生学講座

研究分担者　西谷直之　岩手医科大学 薬学部 臨床薬学講座 情報薬科学分野
米倉佑貴　聖路加国際大学大学院 看護学研究科 看護情報学
林　邦好　京都女子大学 データサイエンス学部

背景

糖尿病や**高血圧**などの生活習慣病は将来の**脳心血管疾患**や**認知症**、**フレイル・要介護**の原因であり、**長期的に継続的な服薬管理が必要**である。

糖尿病 vs. 非糖尿病

脳心血管疾患 **1.6 倍**

(Hirakawa Y, et al. J Epidemiol 2017;27:123–129)

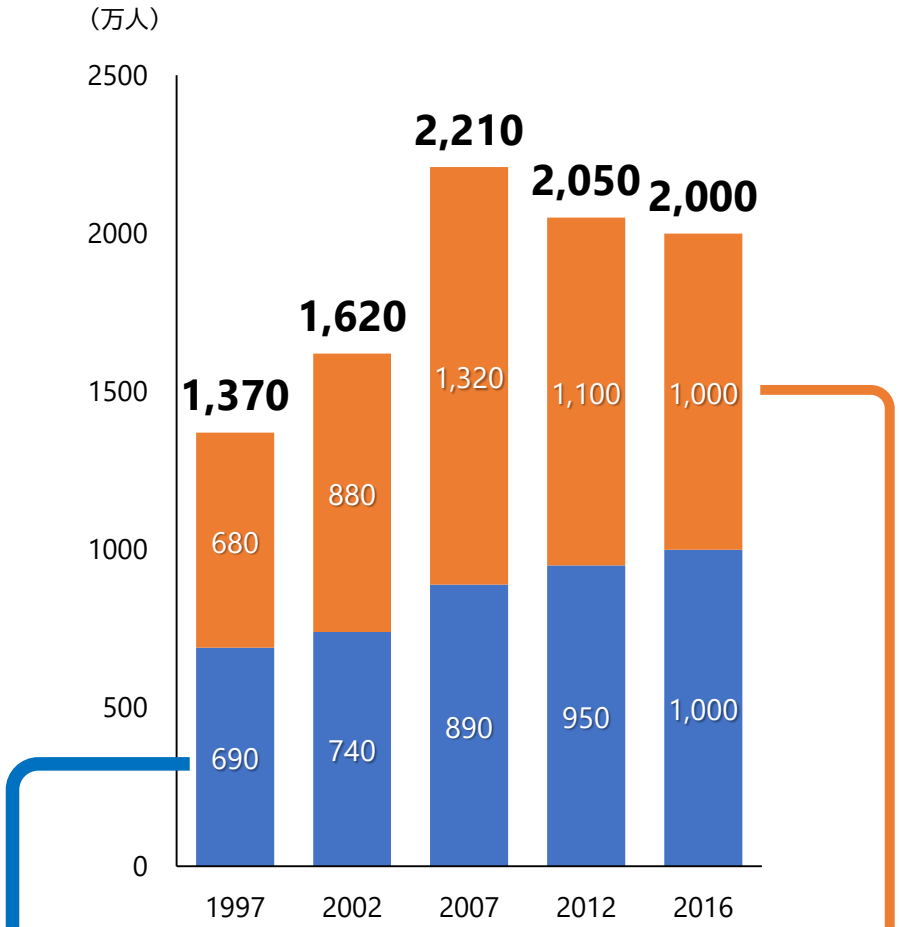
認知症 **1.7 倍**

(Ohara T, et al. Neurology 2011;77:1126–1134)

日常生活動作低下 **1.7 倍**

(Hoang PTN, et al. J Diabetes Investig 2022; 13: 1897–1904)

糖尿病受診中断の実態と疫学



「糖尿病が強く疑われる者」と「糖尿病の可能性を否定できない者」の推計人数の推移

（H28国民健康・栄養調査）

■ 治療中断者割合

- H24国民健康・栄養調査

13.5%

男性　**13.9%**

女性　**13.1%**

治療中断者：治療をこれまで受けたことがある者のうち、「過去に受けたことがあるが現在は受けていない」と回答した者

- J-DOIT2（糖尿病受診中断対策包括ガイド）

8.16 – 8.25%

治療中断者：次回予定受診日から2か月の間受診がない者

本研究から期待される効果

- 従来の治療中断や治療コントロール不良への対策は、中断やコントロール不良が確認されてから行われることが多く、病状が悪化してから治療が再開されるケースが多い。

- 治療中断や治療コントロール不良の予測モデル構築により、治療中断、治療コントロール不良になる前にハイリスク者を早期に発見し、適切に介入することができるようになることが期待される。

目的

- 本研究の目的は、**機械学習**を用いて**生活習慣病（糖尿病及び高血圧）**の

  ① 治療中断予測モデル

  ② 治療コントロール不良予測モデル

  を構築することである。

- **2023年度の研究実施状況**について報告する。

2023年度の研究実施状況

- レセプト情報を用いた治療中断モデルの検討
- 健診情報を用いたコントロール不良モデルの検討

対象

2015–2020年度の岩手支部レセプトデータ
糖尿病　230,044件

糖尿病治療者　843人

2015年度の岩手支部レセプトデータ
傷病名：糖尿病
処方薬：糖尿病治療薬

DPP-4阻害薬
αグルコシダーゼ阻害薬
SU剤
ビグアナイド
チアゾリジン
グリニド
GLP-1受容体作動薬
インスリン製剤
SGLT2阻害薬
糖尿病配合剤

糖尿病治療中断者　166人

2015年度の岩手支部レセプトデータ
受診間隔≥3か月

2015–2020年度の岩手支部レセプトデータ
高血圧症　653,116件

高血圧症治療者　3,570人

2015年度の岩手支部レセプトデータ
傷病名：高血圧症
処方薬：降圧薬

ARB
ACE阻害薬
Ca拮抗薬
利尿剤
βブロッカー
α・αβブロッカー
降圧薬配合剤
降圧薬その他

高血圧症治療中断者　998人

2015年度の岩手支部レセプトデータ
受診間隔≥3か月

統計解析

- アウトカム：糖尿病治療中断者、及び、高血圧症治療中断者
- Classification and Regression Trees (CART) モデルによって説明変数の変数重要度を決定
- ロジスティック回帰モデルを作成

  ※岩手県以外の東北地方のエリアのデータを検証用データとして想定したため、今回はすべてのデータを学習データとしてモデル構築を行った。
- ロジスティック回帰モデルにおける感度、特異度、Area Under Curve (95% confidence interval) [AUC (95% CI)] を計算

CARTモデルによる変数重要度

説明変数	変数重要度
DPP-4阻害薬処方量/年度	34.323
ビグアナイド処方量/年度	22.059
インスリン製剤処方量/年度	18.149
業種	17.187
年齢	11.794
SU剤処方量/年度	11.547
DPP-4阻害薬処方の有無	9.711
チアゾリジン処方量/年度	6.590
ビグアナイド処方の有無	6.477
SU剤処方の有無	6.249
薬処方種類数	5.913
インスリン製剤処方の有無	5.184
標準報酬月額（千円）	4.742
チアゾリジン処方の有無	3.323
GLP-1受容体作動薬処方の有無	1.612
性別	1.372
眼科レセプト発生	1.249
糖尿病配合剤処方の有無	0.874
GLP-1受容体作動薬処方量/年度	0.858
SGLT2阻害薬処方量/年度	0.644
糖尿病配合剤処方量/年度	0.579
栄養指導料算定回数/年度	0.462
αグルコシダーゼ阻害薬処方量/年度	0.393
SGLT2阻害薬処方の有無	0.310
歯科レセプト発生	0.163
雇用/個人事業主	0.163
グリニド処方の有無	0.163
グリニド処方量/年度	0.125

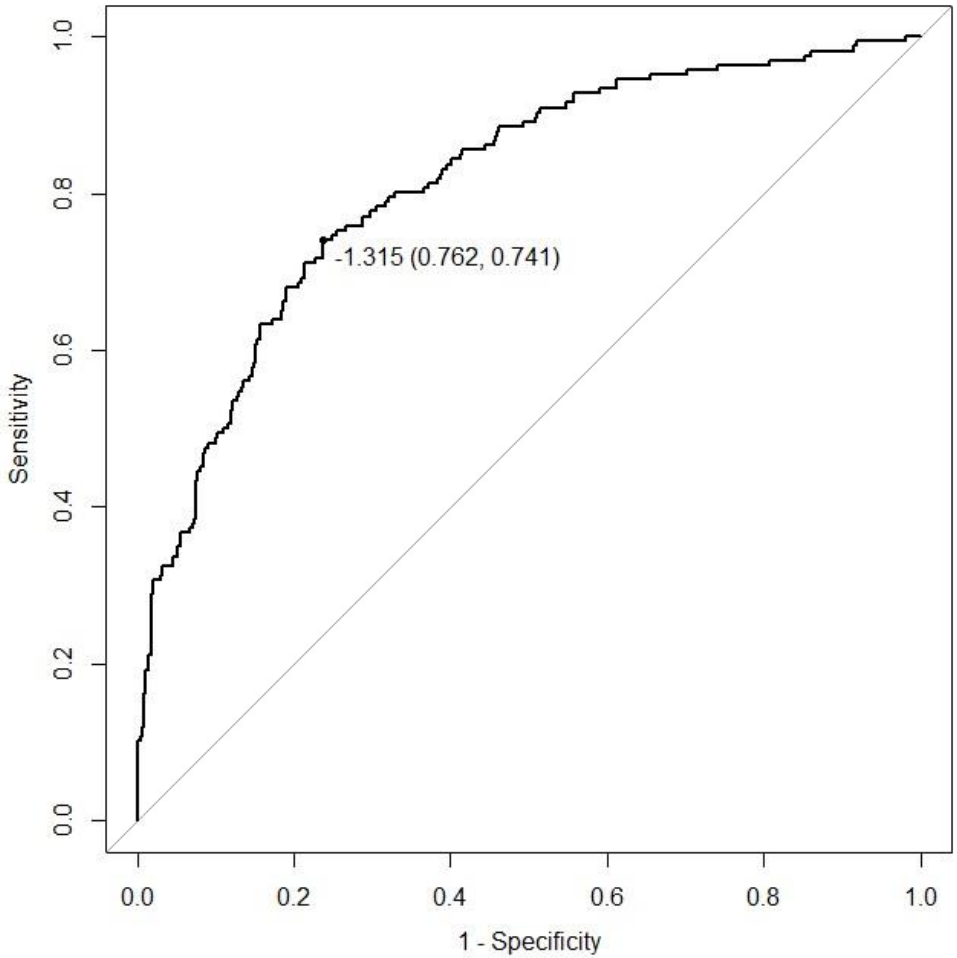
ロジスティック回帰モデルの結果

説明変数		OR	95% CI		
			下限	上限	
年齢		**0.9789**	**0.9587**	**0.9995**	*****
ビグアナイド処方量/年度		**0.9972**	**0.9961**	**0.9984**	*******
DPP-4阻害薬処方量/年度		**0.9956**	**0.9939**	**0.9973**	*******
SU剤処方量/年度		**0.9973**	**0.9953**	**0.9992**	******
インスリン製剤処方量/年度		**0.9451**	**0.9200**	**0.9709**	*******
ビグアナイド処方あり	**vs. なし**	**3.0236**	**1.4845**	**6.1585**	******
DPP-4阻害薬処方あり	**vs. なし**	**3.6652**	**1.9842**	**6.7706**	*******
SU剤処方あり	**vs. なし**	**2.4424**	**1.2110**	**4.9258**	*****
インスリン製剤処方あり	**vs. なし**	**5.7630**	**2.8808**	**11.5287**	*******

***, $P<0.001$; **, $P<0.01$; *, $P<0.05$; . $P<0.1$. OR, odds ratio; CI, confidence interval.

糖尿病治療中断予測モデルのROC曲線

感度：0.741、特異度：0.762、AUC (95%CI) ：0.813 (0.776 – 0.850)



CARTモデルによる変数重要度

説明変数	変数重要度
Ca拮抗薬処方量/年度	261.130
ARB処方量/年度	202.059
降圧薬配合剤処方量/年度	135.506
降圧薬配合剤処方の有無	119.914
Ca拮抗薬処方の有無	117.708
薬処方種類数	63.314
ACE阻害薬処方量/年度	54.277
業種	52.749
ARBの処方の有無	40.896
標準報酬月額（千円）	32.075
利尿剤処方量/年度	29.185
年齢	28.693
α・αβブロッカー処方量/年度	13.751
ACE阻害薬処方の有無	8.054
利尿剤処方の有無	6.889
性別	4.417
降圧薬その他処方量/年度	3.831
α・αβブロッカー処方の有無	2.424
βブロッカー処方の有無	1.850
降圧薬その他処方の有無	1.553
βブロッカー処方量/年度	1.194
雇用/個人事業主	0.605

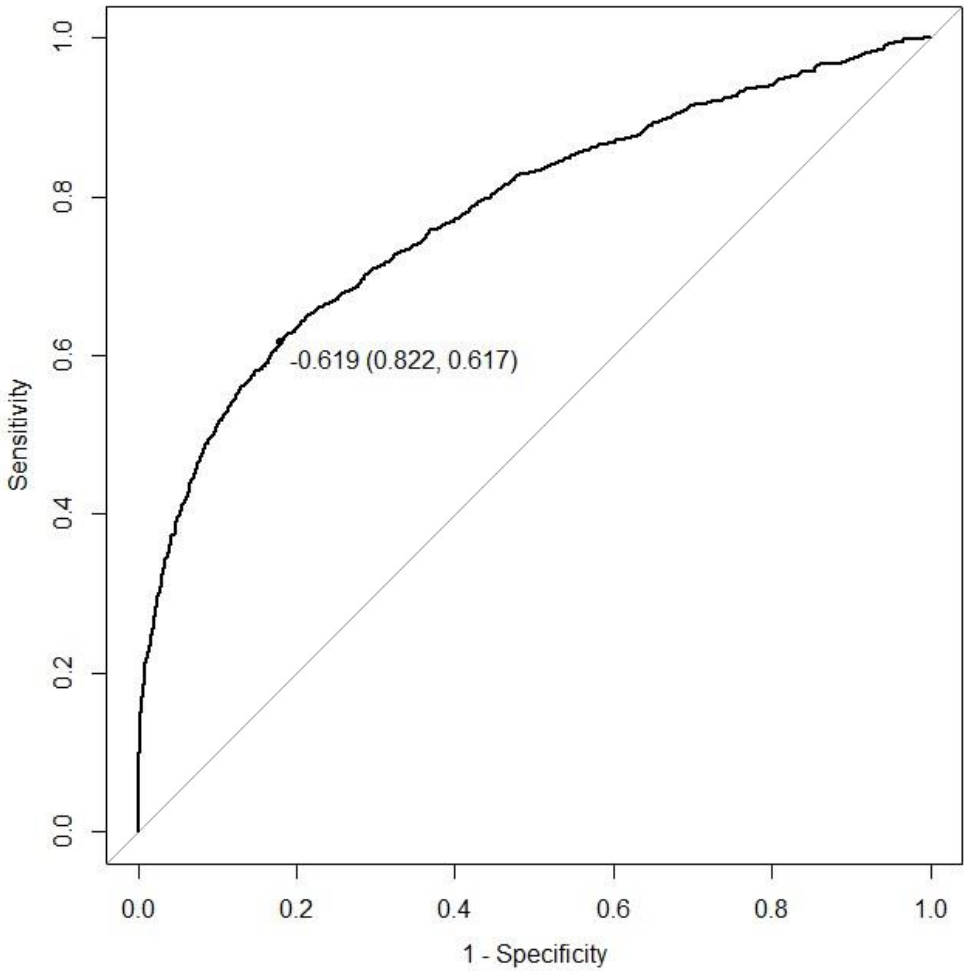
ロジスティック回帰モデルの結果

説明変数		OR	95% CI		
			下限	上限	
年齢		**0.9890**	**0.9791**	**0.9989**	*
Ca拮抗薬処方量/年度		**0.9978**	**0.9973**	**0.9983**	***
降圧薬配合剤処方量/年度		**0.9965**	**0.9951**	**0.9979**	***
Ca拮抗薬処方あり	**vs. なし**	**1.6336**	**1.2446**	**2.1443**	***
ACE阻害薬処方あり	**vs. なし**	**1.5316**	**1.0353**	**2.2658**	*
利尿剤処方あり	**vs. なし**	**3.1734**	**2.0118**	**5.0058**	***
βブロッカー処方あり	**vs. なし**	**6.3804**	**4.5853**	**8.8784**	***
α・αβブロッカー処方あり	**vs. なし**	**1.6943**	**1.0001**	**2.8707**	*
降圧薬配合剤処方あり	**vs. なし**	**1.6906**	**1.0368**	**2.7569**	*
降圧薬その他処方あり	**vs. なし**	**10.6576**	**4.3899**	**25.8743**	***
薬処方種類数		**0.8923**	**0.8736**	**0.9115**	***
飲食・宿泊		**1.9040**	**1.0529**	**3.4430**	*
標準報酬月額（千円）		**1.0008**	**1.0004**	**1.0012**	***

***, $P<0.001$; **, $P<0.01$; *, $P<0.05$; . $P<0.1$. OR, odds ratio; CI, confidence interval.

高血圧症治療中断予測モデルのROC曲線

感度：0.617、特異度：0.822、AUC (95%CI) ： 0.778 (0.760 – 0.796)



治療中断モデルの結果のまとめ

- 治療薬の処方ありと中断リスク：**正**の関連
- 処方量と中断リスクの間：**負**の関連
- 薬処方種類数と中断リスク：に**負**の関連

- 処方量少、薬処方種類少で中断が発生しやすい
  - より**軽症者**でより治療中断が起こりやすい可能性
  - 一方、中断したために1年間の処方量が少なかった可能性

- 診療ガイドラインで第一選択薬より治療困難時に選択される薬剤処方のほうで効果量が大きい傾向
  - **治療困難者**で治療中断がより起こりやすい

2023年度の研究実施状況

- レセプト情報を用いた治療中断モデルの検討
- **健診情報を用いたコントロール不良モデルの検討**

対象

2015–2020年度の岩手支部健診データ
408,397件

糖尿病治療者　1,780人

2015年度の岩手支部健診データ
標準的な質問票
糖尿病治療薬の使用あり

糖尿病治療コントロール不良者
1,540人

2015 – 2020年度の岩手支部健診データ
計測時点のどこかで
空腹時血糖≥130 mg/dL
随時血糖値≥180 mg/dL
HbA1c≥7.0%

高血圧症治療者　4,392人

2015年度の岩手支部健診データ
標準的な質問票
降圧薬の使用あり

高血圧症治療コントロール不良者
2,066人

2015 – 2020年度の岩手支部健診データ
計測時点のどこかで
収縮期血圧≥140 mmHg
拡張期血圧≥90 mmHg

統計解析

- アウトカム：糖尿病治療コントロール不良者、及び、
  高血圧症治療コントロール不良者
- CARTモデルによって説明変数の変数重要度を決定
- ロジスティック回帰モデルを作成
- ロジスティック回帰モデルにおける感度、特異度、AUC (95% CI)を計算

CARTモデルによる変数重要度

説明変数	変数重要度
GGT	21.543
総コレステロール	17.305
AST	16.953
年齢	15.931
LDLコレステロール	14.718
収縮期血圧	13.773
拡張期血圧	13.575
BMI	13.390
ALT	10.703
クレアチニン	8.168
腹囲	7.425
赤血球数	5.582
ヘマトクリット	4.062
同年齢の同性と比べて歩く速度が速い	4.046
HDLコレステロール	3.796
1年間で±3kgの体重変化	3.400
生活習慣改善の意思	2.139
飲酒頻度	1.741
食べる速度	1.590
朝食欠食が週3回以上	1.412
脂質異常症治療薬の服用	1.192
性別	0.871
夕食後に間食（3食以外の夜食）をとることが週に3回以上ある。	0.593
貧血の既往	0.482

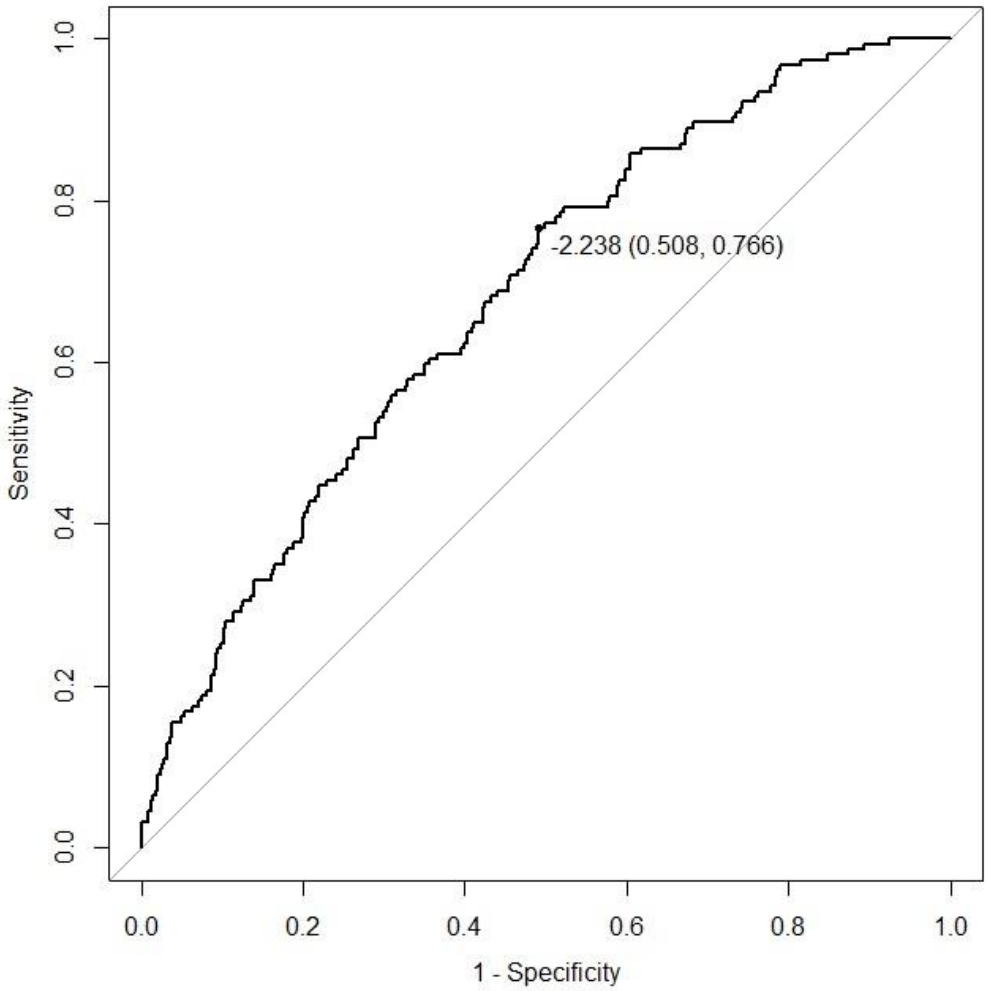
ロジスティック回帰モデルの結果

説明変数		OR	95% CI		
			下限	上限	
GGT		**1.0064**	**1.0002**	**1.0126**	*
LDLコレステロール		**1.0088**	**1.0022**	**1.0154**	**
クレアチニン		**0.5038**	**0.3031**	**0.8374**	**
飲酒頻度：時々	**vs. 毎日**	**1.6969**	**1.0050**	**2.8649**	*
飲酒頻度：ほとんど飲まない（飲めない）	vs. 毎日	0.8728	0.5372	1.4181	

***, *P*<0.001; **, *P*<0.01; *, *P*<0.05; . *P*<0.1. OR, odds ratio; CI, confidence interval.

糖尿病治療コントロール不良予測モデルのROC曲線

感度：0.766、特異度：0.508、AUC (95%CI) ： 0.679 (0.637–0.722)



CARTモデルによる変数重要度

説明変数	変数重要度
収縮期血圧	787.897
拡張期血圧	406.298
腹囲	73.864
BMI	73.820
ALT	51.518
HDLコレステロール	49.862
GGT	47.862
総コレステロール	45.572
AST	43.798
赤血球数	41.260
ヘマトクリット	38.725
年齢	34.665
LDLコレステロール	27.086
クレアチニン	23.516
1日当たりの飲酒量	22.174
飲酒頻度	17.426
空腹時血糖	12.820
生活習慣改善の意思	12.393
性別	11.810
食べる速度	9.379
採血の有無	8.077
HbA1c	7.620
貧血の既往	6.690
血糖降下剤又はインスリンの服用	5.939
20歳時より10kg体重増加	4.702

説明変数	変数重要度
就寝前2時間以内の夕食が週3回以上	3.512
脳卒中の既往	3.455
睡眠で休養が十分とれている	3.394
保健指導利用の希望	2.441
夕食後に間食（3食以外の夜食）をとることが週に3回以上ある。	1.900
心臓病の既往	1.323
脂質異常症治療薬の服用	1.061
現在喫煙	0.514
朝食欠食が週3回以上	0.295
運動習慣（1回30分以上・週2回以上を1年間実施）	0.046

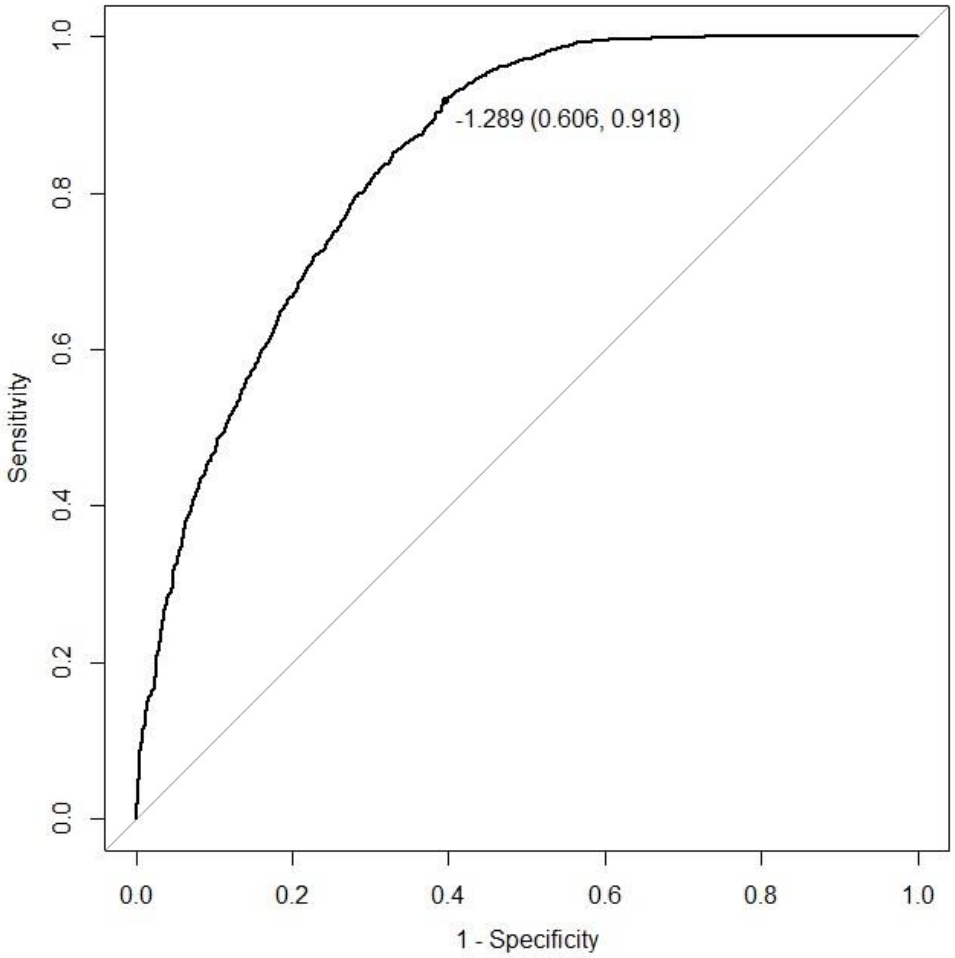
ロジスティック回帰モデルの結果

説明変数		OR	95% CI		
			下限	上限	
収縮期血圧		**1.0894**	**1.0800**	**1.0989**	***
拡張期血圧		**1.0504**	**1.0380**	**1.0629**	***
BMI		**1.0720**	**1.0243**	**1.1219**	**
HDLコレステロール		**1.0074**	**1.0018**	**1.0131**	**
赤血球数		**0.9962**	**0.9926**	**0.9997**	*
ヘマトクリット		**1.0713**	**1.0267**	**1.1178**	**
年齢		**1.0155**	**1.0036**	**1.0276**	*
クレアチニン		**1.9219**	**1.3175**	**2.8037**	***
20歳時より10kg体重増加なし	**vs. あり**	**1.2866**	**1.0706**	**1.5462**	**
脳卒中の既往なし	**vs. あり**	**0.5000**	**0.3277**	**0.7630**	**

***, $P<0.001$; **, $P<0.01$; *, $P<0.05$; . $P<0.1$. OR, odds ratio; CI, confidence interval.

高血圧症治療コントロール不良予測モデルのROC曲線

感度：0.918、特異度：0.606、AUC (95%CI) ： 0.842 (0.831–0.853)



治療コントロール不良モデルの結果のまとめ

- 糖尿病治療コントロール不良モデル
  - クレアチニンと治療コントロール不良：**負**の関連
  - 飲酒頻度が時々（vs. 毎日）：**正**の関連

- 高血圧症治療コントロール不良モデル
  - 20歳時より10kg体重増加なし（vs. あり）：**正**の関連

まとめ

- 岩手支部のデータを用いて、治療中断、及び、治療コントロール不良の判別に重要な予測因子を探索することができたが、予測能は必ずしも高くなかった。

◆ 今後の課題

- 対象者の限定
- 投入する変数の妥当性の検討
- 層化解析
- 感度分析
- 全国データを用いた分析の実施