

# 心電図機械学習を用いた糖尿病スクリーニング手法の検討

植田 蓮 【 コミュニケーション&コラボレーション研究室 】

## ECG-Based Machine Learning Approach for Diabetes Screening

UETA, Ren 【 Communication and Collaboration Lab. 】

### 1 はじめに

糖尿病は世界的に増加しており、2021年の成人有病率は10.5%に達する[1]。日本においても推定1,100万人が罹患し、そのうち約半数が未診断とされる。糖尿病は心血管疾患の主要因であり、発症早期から心筋の電気生理学的特性に変化を来し得ることが報告されており、早期発見と早期介入が重要である。

糖尿病の標準的な診断およびスクリーニングは血液検査に基づくが、採血を伴う侵襲性に加え、受診行動や継続受診の制約により取りこぼしが生じ得る。実際に、定期健診において血液検査が提供されても未受診者が一定割合存在することが報告されており、低侵襲で実装可能性の高いスクリーニング手段の整備が求められる。

近年、糖尿病と心血管疾患との関連性に着目し、心電図を用いた糖尿病関連リスク評価の研究が報告されている。Kimらは心電図と年齢・性別を入力とする深層学習により糖尿病リスクを算出し、その値が非糖尿病患者の将来発症リスクと関連することを示した[2]。一方で、糖尿病は年齢や性別と強く関連する疾患であり、ECG波形そのものに由来する情報の寄与や予測根拠は十分に明らかでない。本研究では、心電図波形の糖尿病バイオマーカーとしての有用性を示し、その予測根拠を生理学的観点から考察する。

### 2 方法

#### 2.1 データ収集

本研究は後ろ向き横断研究であり、2000年4月から2024年2月までに収集された健診データを解析対象とした。心電図は標準12誘導で取得され、サンプリング周波数500Hz、記録長10秒であった。記録時には心電計により0.04-100Hzのバンドパスフィルタが適用されている。

#### 2.2 解析アプローチとデータセット

本研究では、心電図波形を説明変数として、3年以内の糖尿病発症の有無を分類を行う1D-CNNおよび一般化加法モデル(GAM)を構築した。GAMの説明変数については、CNNの畳み込み層が微分や積分に相当するフィルタ特性を獲得しやすいという経験的知見に基づき、従来の振幅指標ではなく、波形の傾きおよび面積を

採用した。目的変数は、データ収集期間の最初の健診受診日から180日以上、3年以内に糖尿病を発症した場合を陽性、3年以上追跡して未発症の場合を陰性と定義した。なお、糖尿病陽性の医学的基準に基づき、 $HbA1c \geq 6.5\%$ 、空腹時血糖  $\geq 126 \text{ mg/dL}$ 、自己申告のいずれかを満たす場合とした。

解析対象は、初めての健診受診時に糖尿病が陰性であり、かつ3年以上の追跡データが存在する受診者とした。受診者1名につき1レコードに限定し、22,513レコード(陽性率4.5%)を抽出した。評価は受診者単位で分割したホールドアウトセットにより実施した。

#### 2.3 CNNによる発症予測モデルの構築

将来発症の予測モデルとして、心電図波形を直接入力とする1次元CNNを構築した。12チャンネル  $\times$  5,000サンプルのテンソルとして、各誘導を別のチャンネルに入力した。各誘導の振幅は学習データの平均と標準偏差に基づいてz-score正規化し、推論時も同一統計量で変換した。

CNNは3層の畳み込み層から構成され、各層のカーネルサイズをそれぞれ257, 129, 65と比較的大きく設定した。これは心電図波形では同様の拍動パターンが時間方向に連続して出現するという特性を考慮し、広い時間文脈を学習させるためである。この設計により、ストライドおよびプーリングを含めた最終畳み込み層の理論受容野は入力長を上回り、波形全体を参照可能となる。その結果、モデルはQRS波やT波といった局所的な形態情報に加え、心拍リズムや全体的な伝導遅延などの大域的な時系列パターンを同時に捉えることができる。

#### 2.4 GAMによる形態特徴量を用いた解析

CNNと同様に3年以内の発症予測を目的変数とし、解釈可能な形態特徴量を用いたGAMを構築した。まず、自動デリネーションにより各心拍のP波、QRS複合体、T波のオンセットおよびオフセットを推定し、波形セグメントを抽出した。各セグメントに10msの移動平均を適用した後、形態特徴量として以下の3指標を算出した。(i) 面積(絶対振幅の時間積分)、(ii) 最大傾き(1次差分の最大値)、(iii) 最小傾き(1次差分の最小値)。特徴量は心拍ごとに計算し、誘導ごとに中央値で集約した。形態特徴量1つと共変量(年齢、性別、

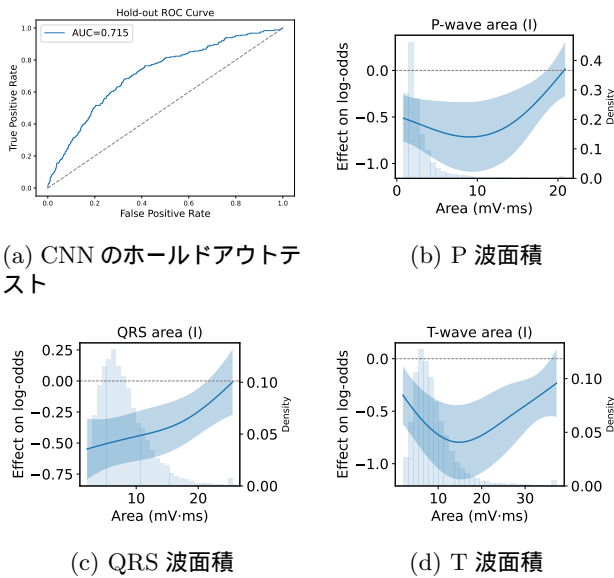


図 1: ROC 曲線と部分依存プロット

BMI, 収縮期血圧) の計 5 変数を説明変数としてモデルを構築し, 推定結果を部分依存プロットとして可視化することで, 各形態特徴量の寄与 (対数オッズ) を評価した。

### 3 結果

#### 3.1 CNN による 3 年以内の糖尿病発症予測

図 1a に CNN のホールドアウトテストにおける ROC 曲線を示す。評価の結果, モデルは AUC 0.715 (95% CI: 0.680–0.746) を達成した。感度 90% を目標とした閾値設定での評価では, 感度 0.85, 特異度 0.38, PPV 0.06, NPV 0.98 を示した。

#### 3.2 GAM による形態特徴量と糖尿病の関連解析

解析には 12 誘導すべてを用いたが, 本稿では I 誘導の主要な結果を示す。図 1b, 1c, 1d に GAM による部分依存プロットを示す。GAM による解析では, P 波面積, QRS 波面積, T 波面積に関して, いずれにおいても面積の増加が糖尿病陽性確率の上昇と関連していた。一方で, それぞれの対数オッズは 0 を下回っており, 特徴量が予測に与える直接の寄与は限定的であった。また, GAM の AUC はいずれも約 0.82 であったが, 本稿では共変量による影響を踏まえ, 形態特徴量と糖尿病発症との関連の解析を主解析とした。

### 4 考察

本研究では, 標準 12 誘導心電図から 3 年以内の糖尿病発症を予測する 1D-CNN を構築するとともに, 予測の根拠を工学的・生理学的観点から検討するため, GAM による解析を行なった。

構築した CNN はホールドアウト評価において AUC 0.715, NPV 0.98 を達成した。AUC が 0.5 を明確に上回ったことは, 臨床的に糖尿病発症前段階の情報が心

電図波形中に含まれ得ることを示唆する。また, 高い NPV は, 本手法が低リスク群を高い信頼度で除外し, 追加検査の優先順位付けを支援する補助的スクリーニングとして有用である可能性を示す。近年, スマートウォッチや家庭用心電計の普及により非侵襲的な心電図取得が一般化しつつあり, 本知見は日常環境での早期発見と早期介入に向けた実装可能性を後押しすると考えられる。

一方, GAM 解析では波形の面積および傾きに基づく形態特徴量を用い, 各セグメントにおける面積の増加が糖尿病のオッズ上昇と関連する傾向を示した。QRS 複合体は心室の脱分極過程を表すものであることから, 糖尿病を発症する前に心室脱分極過程に由来する形態情報が心電図波形中に含まれ得ることを示唆する。

信号処理の観点において, CNN の畳み込み層が担うフィルタリング処理は, 入力信号に対する微分 (勾配検出) や積分 (平滑化・エネルギー抽出) といった演算特性を内包する。本研究で GAM の入力として設計した傾きおよび面積は, CNN がブラックボックス的な学習過程で獲得し得るこれらの潜在表現を, 解釈可能な物理量として明示的に近似したものと位置づけられる。GAM において積分指標である面積の有用性が示されたことは, 1D-CNN が学習過程において積分的なフィルタ特性を通じて波形のエネルギー量を捉え, 心室脱分極の異常という生理学的実体に基づいて予測を行っている可能性を示唆する。

### 5 まとめ

本研究では, 標準 12 誘導心電図から糖尿病発症リスクを予測する 1D-CNN を構築し, AUC 0.715 および NPV 0.98 というスクリーニングに適した性能を得た。さらに, GAM を用いた解析によって, 深層学習モデルが獲得した表現が, 波形の面積や傾きといった解釈可能な特徴と整合することを確認し, モデルの生理学的妥当性および説明性を担保した。

実臨床への導入に向けては, 対象集団の有病率や検査目的に応じた適切な閾値設定の検討が不可欠である。今後は, 多様な患者背景を含む外部データでの検証に加え, 実環境下での有用性を評価する前向き研究を実施し, 予防医療における早期発見・介入支援システムとしての確立を目指す。

### 参考文献

- [1] International Diabetes Federation. IDF Diabetes Atlas, 10th edn. Brussels, Belgium, 2021.
- [2] Kim J, Yang HL, Kim SH, *et al.* Deep learning-based long-term risk evaluation of incident type 2 diabetes using electrocardiogram in a non-diabetic population: a retrospective, multicentre study. *EClinicalMedicine*. 2024;68:102445. doi:10.1016/j.eclinm.2024.102445.