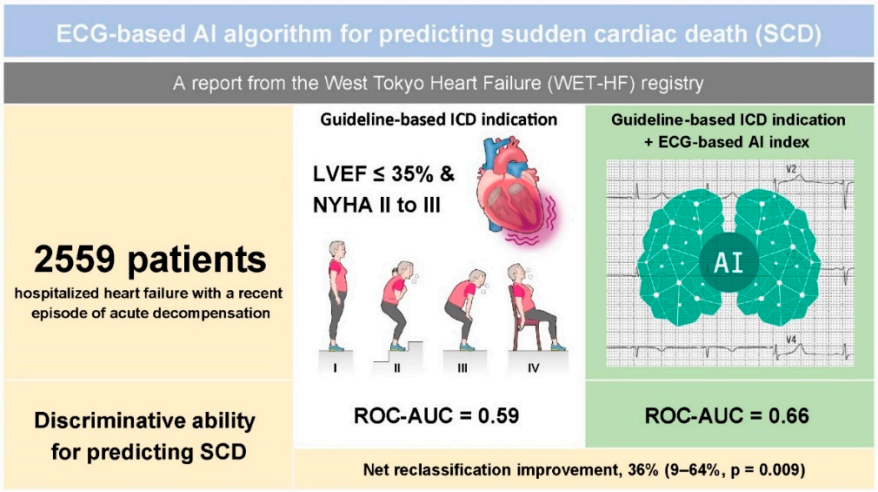
205　画像情報と深層学習を応用した心疾患患者の予後予測	白石　泰之

【目的】心臓突然死（Sudden Cardiac Death：SCD）は心不全患者における主要な死因の一つであり、その予測と適切な介入（例：植込み型除細動器［ICD］）が臨床上極めて重要である。しかしながら、既存の SCD 予測アルゴリズムの精度は十分ではなく、その有用性は限定的である。当研究では、心電図（ECG）データをもとに人工知能（AI）を使用した新しい SCD 予測指標（ECG-AI index）を開発し、さらに ECG-AI index と従来の予測指標（既存のガイドライン推奨）を組み合わせることでその精度が向上するか検証を行う。

【方法】東京首都圏の基幹病院から構成される、多施設共同心不全レジストリ（West Tokyo Heart Failure Registry）に登録された 2,559 名の心不全入院患者を対象とした。入院中・退院前に記録した心電図（各誘導における時系列電位データ）とアウトカム指標（SCD の発生）を AI モデルに学習させ、各患者における SCD 予測率（＝ECG-AI index）を算出した。患者は ECG-AI index を作成するコホート（Derivation：病院 A）、校正するコホート（Validation：病院 B）、そして最終的に精度を確認するコホート（Test：病院 C と D）にそれぞれ分割された。従来の SCD 予測指標である左室駆出率（LVEF）と NYHA 機能分類（心不全の重症度分類の一つ）、ならびに競合リスク（非心臓突然死［non-SCD］）を考慮した上で、ECG-AI index とエンドポイントとの関連性を、回帰モデルを用いて評価した。主要エンドポイントは観察期間中の SCD 発生ならびに ICD 作動（電気的除細動と抗頻拍ペーシング）とした。

【結果】各コホートの患者背景は、年齢 73～78 歳また LVEF 40～48%であり、NYHA 機能分類Ⅱ～Ⅲの割合はそれぞれ 75.2%、92.4%、93.4%であった。また ICD 植込みは 3.0～8.4%の患者で観察された。最終 Test コホートにおいて、236 例（1,077 例中の 21.9%）の死亡が発生し、その内訳は 48 例の SCD と 188 例の non-SCD となり、さらに 4 例の ICD 作動が観察された。従来の予測指標（LVEF≦35%と NYHA 機能分類Ⅱ～Ⅲ）に ECG-AI index を追加することで、SCD 予測能は有意に改善した（ROC-AUC、0.66 vs. 0.59［p=0.017 for Delong's test］; p=0.11 for Hosmer-Lemeshow test［Calibration］; net reclassification improvement、36%［95%信頼区間 9～64%］、p=0.009）。競合リスク（non-SCD）を考慮した Fine-Gray 競合モデルにおいても、ECG-AI index は有意に SCD 発生と関連していた（調整後 subdistributional hazard ratio、1.25［95%信頼区間 1.04～1.49］; p=0.015）。さらに、ECG-AI index が高値になるほど、死因に占める SCD の割合（proportional risk）は有意に上昇していることが観察された（ECG-AI index の三分位で分割：低リスク群 16.7%、中等度リスク群 18.5%、高リスク群 28.7%［p for trend=0.023］）。最後に、ECG-AI index を重要な患者背景による層別解析で評価したところ、年齢 75 歳以下、非虚血性心不全、LVEF が 35%上回る患者群においても同様の精度が確認された（Shiraishi Y. doi: https://doi.org/10.1101/2022.03.20.22272659）。

Central illustration of the present study



205. 画像情報と深層学習を応用した心疾患患者の予後予測

白石　泰之

慶應義塾大学 医学部 難治性循環器疾患病態学寄附講座

Key words：心臓突然死，心電図，人工知能，心不全，競合リスク

緒　言

心不全は深刻なQOL低下、ならびに高い死亡率を伴う重大な臨床症候群であり、心不全患者のおよそ半数は診断後5年以内に死亡するとされている［1］。心不全の罹患率はアジア（1.2～6.7%）や欧米諸国（1～14%）で高く、人口の高齢化に伴って今後も増加が予想され、世界的な公衆衛生上の大きな問題となっている。中でも、致死的な不整脈に起因することの多い、心臓突然死（Sudden Cardiac Death：SCD）は心不全患者における心血管死亡の約50%を占めるという報告もあり、SCDを予防するための植込み型除細動器（Implantable Cardioverter-Defibrillator：ICD）が一般に使用されている。しかしながら、ICDの植込み手術は侵襲的であり、さらにICDを植込まれた患者の約50%が不適切な電気的除細動（治療適応のない不整脈でのショック作動）を経験し、これらはQOLを大幅に低下させ、うつ病等の発症リスクを上げ、死亡率まで上昇させることが知られている。したがって、個々の心不全患者においてSCDの発生リスクを正確に評価することは、適切なICD使用に際して、医療者や患者の意思決定に極めて重要である。

現在のSCDリスクの評価方法は、主に左室駆出率（Left Ventricular Ejection Fraction：LVEF≦35%）ならびに心不全患者の重症度指標の一つであるNew York Heart Failure（NYHA）機能分類に基づいている。しかしながら、この評価方法の精度が低いことが近年報告され［2］、結果として臨床現場ではICDの過剰ないし過少使用につながっている。さらに、LVEFが高度に低下していない心不全患者においてもSCDが頻繁に発生している事実があるにも関わらず、LVEF 35%を上回る患者に対してSCDリスクを層別化するアルゴリズムが現状存在しないことも問題である。

人工知能（Artificial Intelligence：AI）は、情報量の多い複雑なデータからある事象の起こる確率を予測するのに有用であり、例えば、畳み込みニューラルネットワーク（Convolutional Neural Network：CNN）を利用した心電図（ECG）AIモデルは、病気の診断［3］、心機能の推定［4］、あるいは予後予測［5］に有用な可能性を示している。そこで我々は心電図AIモデルにより、SCDリスクを分類するための重要な特徴を検出し、リスク層別化を改善することで、医療資源のより適切な利用につながると仮定した。当研究では、前向き観察研究データベースを用いて、心不全患者のSCDを予測する心電図AIモデル（以下、ECG-AI index）を作成し、その精度を評価することを目的とした。

方　法

1. データ収集

West Tokyo Heart Failure Registry（WET-HF）は、東京首都圏の6つの三次医療機関（慶應義塾大学病院、杏林大学病院、榊原記念病院、聖路加国際病院、東京都済生会中央病院、埼玉医科大学国際医療センター）が参加している前向き心不全コホートである［6］。各施設の専属の臨床研究コーディネーターの支援のもと、急性増悪のために緊急治療を要した心不全入院患者の連続登録を原則としている。東京都済生会中央病院と埼玉医科大学国際医療センターの2施設を除く、4施設の患者の12誘導心電図データを各施設の電子カルテシステムから抽出した。心電図データは、500 Hzのサンプリングレートで全12誘導の2.5秒間の時系列電位データが記録された（V1・II・V5誘導は10秒間のデータあり）。同一患者で複数の心電図記録がある場合には、最も退院に近いデータを解析に使用した。2006年～2017年までに最終的に生存退院した2,559名の心不全患者を対象に、ECG-AI indexの開発・評価のために、登録さ

れた医療機関に応じて 3 つの独立したコホートに患者を分類した（Derivation、Validation、Test cohorts（図 1））。

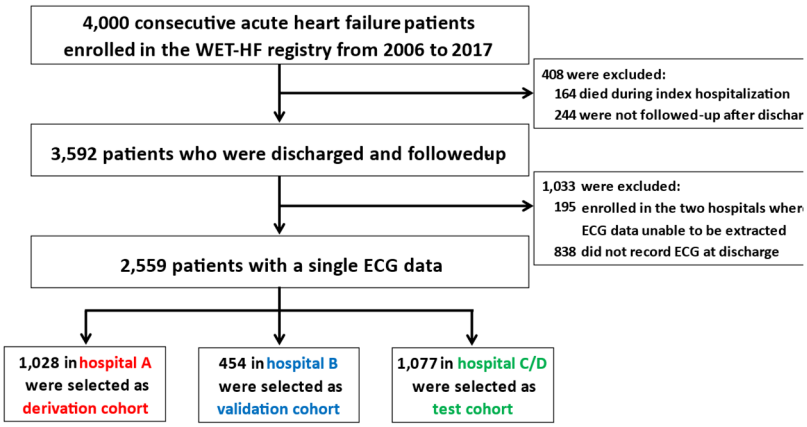


図 1. 登録患者フローチャート

2. モデル作成

CNN と Long Short-Term Memory（LSTM：recurrent neural network［RNN］の一種）を使用して SCD 予測モデルを作成した（図 2）。このモデルは、比較的低い計算コストで「形状パターン」を特徴づけるのに適した 1 次元 CNN を 4 層と、比較的高い計算コストを要する時系列データ解析に適した LSTM を 2 層重ねたニューラルネットワークで構成された。初期学習率 0.0001 で RMSProp オプティマイザーを用いて、binary cross-entropy 損失を最小化するようにモデルの学習を実施した（Derivation）。モデルは各学習（epoch）終了時に Validation データセットを使用して計算し、すべての学習終了時に最も精度の高いモデルを選択し、最終的に Test cohort において一度だけその精度を計算した。モデルの学習には Tensorflow framework version 2.2.0、Python version 3.6.8 を使用した。

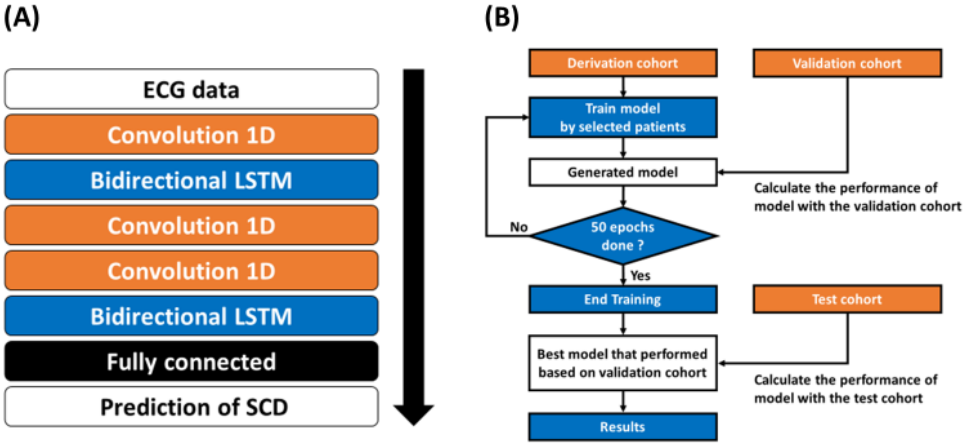


図 2. AI モデルの構造（A）と AI モデル作成の各プロセス（B）

ニューラルネットワークモデルは 1 次元 CNN および LSTM で構築され、全結合層（Fully connected）を通じて、最終的に SCD の予測確率として出力される（図 2A）。被験者は 3 つのコホートに分割され、Early stopping を用いて最良のモデルを作成後、モデル作成に関与していない Test cohort でそのモデル精度を確認した（図 2B）。

3. 評価項目

当解析の評価項目は SCD と ICD 作動（電気的除細動と抗頻拍ペーシング）の複合エンドポイントとした。当研究に

おいてはSCDを正確に評価するために、死亡イベントは研究組織内のイベント判定委員会で評価された。すべての死亡イベントは研究分担医師が確認し、米国心臓協会・心臓病学会（AHA/ACC）のステートメントに基づいて死亡形態を分類し、最終的にイベント判定委員会が死因を決定した。SCDは上述のAHA/ACCのステートメントより、それまで安定していた患者での予期せぬ原因不明の死亡、または心血管系以外の原因が明らかでない不整脈によると思われる死亡（昏睡状態から蘇生を試みた後に死亡した患者を含む）と定義された。ただし、死亡した患者で24時間以上にわたって目撃者のいない場合は「不詳の死亡」と定義し、当研究ではSCDとは明確に区別した。ICD作動は定期外来受診時、あるいは不整脈エピソードやSCDが疑われる場合にデバイスチェックを実施して確定した。

4. **統計解析**

複合SCDイベントの予測に関して、従来の診療ガイドラインのICD推奨事項（LVEF≦35%およびNYHA機能分類Ⅱ～Ⅲ）、またはECG-AI indexとの組み合わせ（ECG-AI複合モデル）の識別能をROC-AUC（ロジスティック回帰分析および pairwise Delong's 検定を使用）、予測確率と実測確率のプロットと Hosmer-Lemeshow 検定によるCalibration性能の評価を行った。次に、事前に計画していた各サブグループ間でのモデル性能の評価を行った（① 年齢：75 歳以上 vs.75 歳未満、②性別：男性 vs.女性、③ 病因：虚血 vs.非虚血）。LVEF 毎（≦35%、35～50%、50%≧［およびカットオフ値を左記以外にも複数設定］）にECG-AI indexの精度、複合SCDイベントとnon-SCD（非心臓突然死）の頻度も評価した。また、競合リスク（SCD vs. non-SCD）を考慮したFine-Grayモデルを使用した生存解析を併せて実施し評価した。

複合SCDイベントとnon-SCDのリスク比を可視化するため、ECG-AI index単独とECG-AI複合モデルのそれぞれによる予測リスクの三分位値で3群に分け、群別のSCDとnon-SCDイベントの発生率を、傾向検定を用いて評価した。すべての解析はR version 4.0.3 (R Foundation for Statistical Computing、Vienna、Austria、2008) のtidyverse、tidymodels、pROC、PredictABEL、survminer、cmprsk、riskRegression、survivalパッケージを使用して実施した。

結果および考察

1. 患者背景

3つのコホートの患者の年齢は73～78歳（中央値）、LVEF 40～48%（中央値）、NYHA機能分類Ⅱ～Ⅲの割合は75.2～93.48%（中央値）、NYHA機能分類Ⅱ～Ⅲの割合は75.2～93.4%であった（each p＜0.001）。ICD使用率は3.0～8.4%であった（each p＜0.001）。ICD 使用率は 3.0～8.4%といずれの施設でも頻度は低かった。Test コホート（n＝1,077）においては、236例（21.9%）が退院後3年間で死亡し、その内訳はSCDが48例（20.3%）でnon-SCDが188例（79.7%）であった。また、適切なICD作動は4例で認められた。

2. モデル性能

ECG-AI index単独での3年間のSCDイベント予測能はROC-AUC＝0.62（95%信頼区間 0.54～0.70）であった。従来のSCD予測の推奨事項（LVEF≦35%とNYHA機能分類ⅡからⅢ）とそこにECG-AI indexを追加したモデル（ECG-AI複合モデル）の性能を比較したところ、ECG-AI複合モデルで予測能は有意に改善した（ROC-AUC：0.66 vs. 0.59［p＝0.017 for Delong's test］、p＝0.11 for Hosmer‐Lemeshow test［Calibration］; net reclassification improvement：36%［95%信頼区間 9～64%］、p＝0.009）（図3）。サブグループ解析では、75歳以下、非虚血性心筋症においてECG-AI複合モデルの精度が優れていた（図4）。また、LVEF別においては LVEF 35～60%の患者群において ECG-AI indexの優れたSCD予測能が観察された。

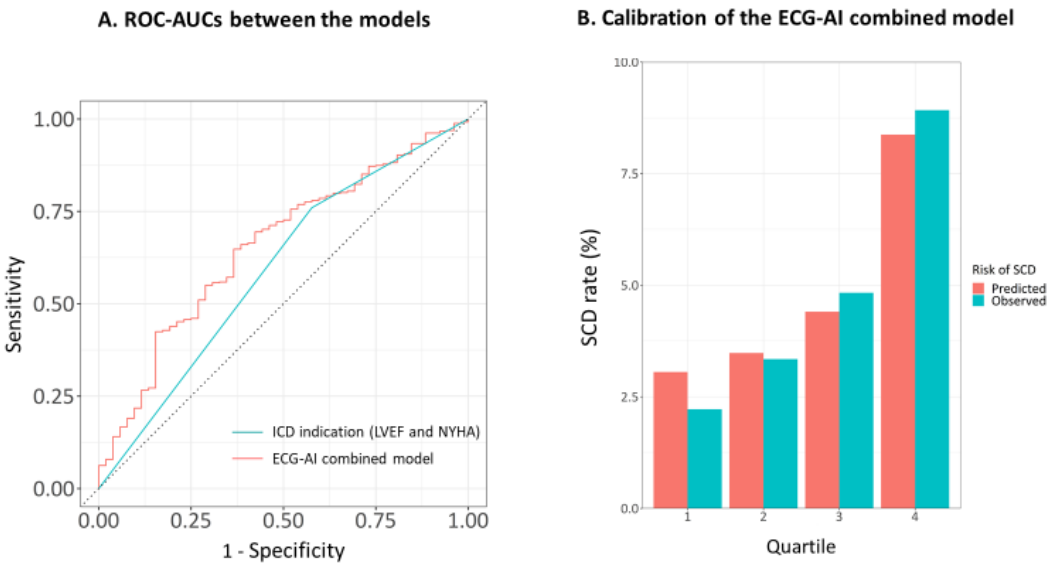


図 3. 各モデルの性能：ROC-AUC（A）と Calibration（B）の 予測値 vs. 実測値）
患者集団を予測 SCD 確率で 4 分割し、各集団での予測値・実測値を比較した（**B**）。

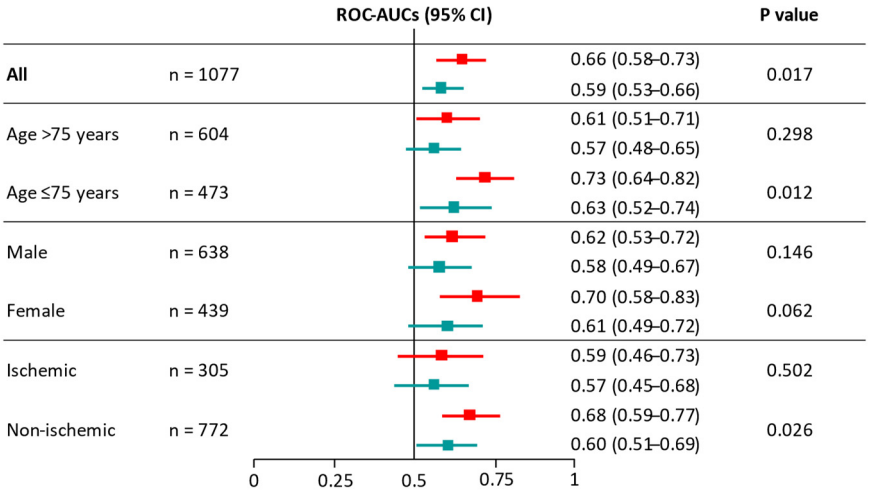


図 4. ECG-AI 複合モデル（緑色）と従来の SCD 予測指標（赤色）のサブグループ間での比較

次に、競合リスク（non-SCD）を考慮した Fine-Gray 競合モデルにおいても、ECG-AI index ならびに従来の SCD 予測指標はともに有意に SCD 発生と関連していた（ECG-AI index に対する調整後 subdistributional hazard ratio：1.25［95%信頼区間 1.04～1.49］、p=0.015；従来指標に対する調整後 subdistributional hazard ratio：1.98［95%信頼区間 1.11～3.54］、p=0.020）。競合リスクモデルにおいて、Bootstrap 法（500 セット）を使用した ECG-AI 複合モデルの精度は Concordance index=0.65（95%信頼区間 0.62～0.69）であった。

最後に、ECG-AI index および ECG-AI 複合モデルの予測リスクに応じて（三分位）、SCD と non-SCD の占める割合を評価した。図 5 に示すように、各モデル推定リスクが高値になるほど、死因に占める SCD の割合は有意に上昇していることが観察された（ECG-AI index：p for trend=0.023、ECG-AI 複合モデル：p for trend<0.001）。

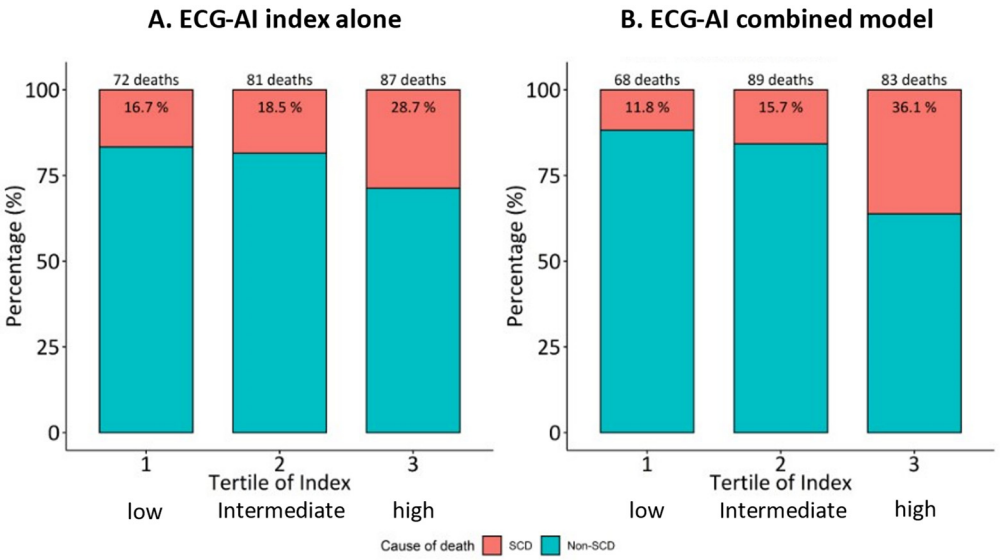


図 5. モデル推定リスク別の SCD と non-SCD の割合

当研究では、従来の SCD 予測アルゴリズムと比較して、我々の開発した心電図 AI を追加することでより正確な SCD 予測が心不全患者において可能となったことを報告した。また、SCD 予測において重要なのは SCD を特異的に予測する（＝non-SCD には関連しにくい）ことであり、同 ECG-AI index が高値になるほど SCD の死因に占める割合が上昇していることが極めて画期的であった。さらに、LVEF が 35%を上回る患者での SCD 予測アルゴリズムがない中で、ECG-AI index の精度が最も良好であったのは LVEF 35～60%であったことも重要な知見と考えられた。

コンピューターやニューラルネットワーク技術の向上は、より深いネットワークパターンの開発を可能にし、例えば、2 次元 CNN を用いた心電図 AI モデルは、年齢と性別を予測し、正常な洞調律の心電図から心機能や潜在的な心房細動の有無を検出することが報告されている［4］。さらに最近、我々は 1 次元 CNN と RNN を組み合わせた心電図 AI モデルが救急現場で緊急心臓カテーテル治療を要する胸痛患者を高い精度で特定できることを報告した［7］。しかしながら、複雑なタスクの中には RNN を必要とするものもあるが、RNN を用いたモデルの性能が CNN よりも優れているかは不明であり、RNN または CNN に適したタスクについてもコンセンサスはまだ得られていない。

SCD に関連する心電図パターンは過去に報告されているが（脱分極や再分極異常、遅延電位など）、個々の識別能は十分でないため臨床現場では重要視されていない。我々の心電図 AI が何を捉えているのか gradient-weighted class activation mapping を用いて調べると、脱分極や再分極異常に関わる部位に着目していることが分かった（データ未提示）。ECG-AI index は SCD に関わる特徴を統合した指標であり、臨床的にも合理的な結果であると考えられた。

当研究では、患者数やイベント数が絶対的に不足していることが一番の限界であり、特にイベント数が増えることでより高精度の SCD 予測が達成される可能性がある。正確な SCD 予測は ICD 植込みに必須であり、今後はより大規模な集団を対象に、この結果が再現できるかを検証していくことが必要である。

共同研究者・謝辞

本研究の共同研究者は、慶應義塾大学医学部循環器内科の後藤信一、新美望、勝俣良紀、佐野元昭、福田恵一、香坂俊、杏林大学医学部循環器内科の合田あゆみ、河野隆志、済生会中央病院循環器内科の武井眞、聖路加国際病院循環器内科の西畑庸介、榊原記念病院循環器内科の佐地真育、吉川勉であり、この場を借りて深く御礼申し上げます。

文　献

1) Virani SS, Alonso A, Aparicio HJ, et al. Heart Disease and Stroke Statistics—2021 Update: A Report From the American Heart Association. Circulation. 2021 Feb 23;143(8):e254-e743. Epub 2021 Jan 27. PMID: 33501848 DOI: 10.1161/CIR.0000000000000950

2) Fukuoka R, Kohno T, Kohsaka S, et al. Prediction of sudden cardiac death in Japanese heart failure patients: international validation of the Seattle Proportional Risk Model. Europace. 2020 Apr 1;22(4):588-597. PMID: 32155253 DOI: 10.1093/europace/euaa002

3) Goto S, Mahara K, Nelson LB, et al. Artificial intelligence-enabled, fully automated detection of cardiac amyloidosis using electrocardiograms and echocardiograms. Nat Commun. 2021 May 11;12(1):2726. PMID: 33976142 DOI: 10.1038/s41467-021-22877-8

4) Attia ZI, Kapa S, Jimenez FL, et al. Screening for cardiac contractile dysfunction using an artificial intelligence–enabled electrocardiogram. Nat Med. 2019 Jan;25(1):70-74. Epub 2019 Jan 7. PMID: 30617318 DOI: 10.1038/s41591-018-0240-2

5) Raghunath S, Cerna AEU, Jing L, et al. Prediction of mortality from 12-lead electrocardiogram voltage data using a deep neural network. Nat Med. 2020 Jun;26(6):886-891. Epub 2020 May 11. PMID: 32393799 DOI: 10.1038/s41591-020-0870-z

6) Shiraishi Y, Kohsaka S, Sato N, et al. 9-Year Trend in the Management of Acute Heart Failure in Japan: A Report From the National Consortium of Acute Heart Failure Registries. J Am Heart Assoc. 2018 Sep 18;7(18):e008687. PMID: 30371201 DOI: 10.1161/JAHA.118.008687

7) Goto S, Kimura M, Katsumata Y, et al. Artificial intelligence to predict needs for urgent revascularization from 12-leads electrocardiography in emergency patients. PLoS One. 2019 Jan 9;14(1):e0210103. PMID: 30625197 DOI: 10.1371/journal.pone.0210103