

高齢慢性心不全患者の大規模データベースの構築と 人工知能を用いた心不全病態の解明

産業医科大学 第2内科学 循環器内科・腎臓内科

稗田 道成

(共同研究者)

九州大学大学院医学系学府医科学専攻	石川 清之助
原土井病院・臨床検査技師	力丸 佳代
原土井病院・臨床検査技師	山本 裕子
原土井病院・臨床検査技師	野口 陽子
原土井病院・リハビリテーション科	樋口 卓篤
原土井病院・リハビリテーション科	西門 美樹
原土井病院・リハビリテーション科	城戸 賢次郎
原土井病院・リハビリテーション科	池江 真弥
原土井病院・看護部	金替 孝子
原土井病院・薬剤部	中道 真理子
原土井病院・歯科	田淵 芳央子
原土井病院・歯科	岩佐 康行
原土井病院・内科	丸山 徹

はじめに

本邦では、高齢化が進み、それに伴い心不全患者数は急速に増加し、2040年頃には140万人にまで達する。心不全は、全ての心血管疾患の終末像であり、加齢に伴い心不全を発症しやすくなる⁽¹⁾。実に、心不全入院患者の50%以上は75歳以上の高齢者であり⁽²⁾、左室拡張障害や再入院率が多いという高齢者特有の病態がある⁽³⁾。この増加する高齢心不全患者への対応は、医学的問題と同時に高齢者特有の社会的・心理的・身体的側面を同時に評価することが極めて重要である。

そこで本研究の目的は、高齢慢性心不全の特徴を把握し、その病態の本質を解明すると共に、効果的な治療管理指針を構築するためのリアルワールドエビデンスを創成することとした。研究計画の通り、高齢慢性心不全患者の医療的側面・社会的側面を反映した高齢者心不全データベース構築を行い、そのデータベースを利用して機械学習による解析を行った。高齢者心不全の未知の病態を明らかにするために、これまで医学研究で最近使用され始めている教師なし学習解析であるK-means法やUMAP (Uniform Manifold Approximation and Projection) を用い、更に最新の解析手法であるPaCMAP (Pairwise Controlled

Manifold Approximation)を使用した。これらを用い、新規集団や特徴量を同定するためにクラスター解析を行った。同定された複数のクラスター集団で生存時間分析をおこなった。

結 果

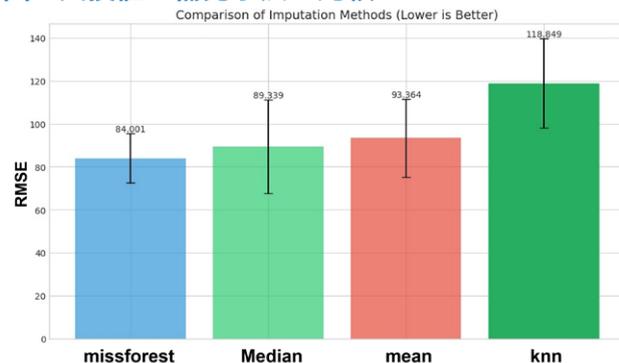
1. 変数選択とデータの性質

高齢者データベース (n=3,475) から778名の心不全患者を抽出し、年齢・性別・BMI・左室駆出率 (LVEF) をMatchingさせた446名を解析対象とした。集計したリアルワールドでの臨床的変数を処理する過程で、まず98の初期変数から欠損率が50～80%と極めて高い変数 (飲酒、Dダイマー、喫煙等) を除外した。これは、過度な補完が解析結果に支配的な影響を及ぼすのを避けるために行った。また、多重共線性解析に基づき、身長・体重をBMIに集約するなど、統計学的な冗長性を排除した上で、臨床的に解釈可能な34変数 (臨床属性、心エコー、血液検査、併存疾患、弁膜症) を選定した。

2. 前処理の最適化：欠損値の補完

解析対象とした34変数のうち23変数に欠損が認められ、特にE/e' (27.8%) やDcT (20.4%) で高率であった。これらの欠損値の補完手法として、中央値、平均値、K-Nearest Neighbors (KNN)、およびRandom Forestベースの反復補完であるmissForestを比較した (図1)。5分割交差検証による評価の結果、missForestによる人工欠損に対する実データと差異指標であるRMSE (二乗平均平方根誤差) は 84.00 ± 11.42 と最も低く、変数間の非線形関係を保持できる点からも本手法を採用した。その後、各変数をStandard Scalerにより平均0、分散1に標準化して対応した。

図1 欠損値と補完手法の比較



変数のなかでE/e' (27.8%)やDcT (20.4%)で欠損率が認められた。複数の補完手法を比較した結果、Random Forestベースの「missforest」が最も低いRMSEであったため、補完手法として採用した。

3. 機械学習モデル (次元削減手法・クラスタリング手法) の選定

クラスタリング手法 (K-means、Ward、Gaussian Mixture Model (GMM)、Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN)) および次元削減手法 (principal component analysis (PCA)、t-Distributed Stochastic Neighbor Embedding (t-SNE)、UMAP、PaCMAP) との組み合わせである計105通り (次元削減の比較候補 PCA: 10次元 (クラスタリング用)、tSNE: 2次元、perplexity=30、max_iter=1000、UMAP: 2次元、min_dist=0.0、metric=euclidean、n_neighbors=15/30/50、PaCMAP:

2次元、MN_ratio=0.5、FP_ratio=2.0、distance=euclidean、n_neighbors=15/30 クラスタリングの比較候補、K-Means : k=3/4/5/6 (n_init=10、max_iter=300)、Ward (階層クラスタリング) : k=3/4/5/6 (linkage=ward) GMM : k=3/4/5/6 (covariance_type=full、n_init=10)、HDBSCAN : min_cluster_size=15/20/30、min_samples=5/10 (cluster_selection_method=eom)、以上をまとめると比較実験数=7(次元削減)×15(クラスタリング)=105通り)を総当たりで検証した(図2)。評価系として、シルエット係数、Davies-Bouldin Index (DBI)、Calinski-Harabasz Index (CHI)の3つの内部指標に加え、「層別化」という目的に反するノイズ(分類不能例)の発生率を制約条件とした。次元削減の評価として、PCAによる主成分分析では、分散の80%を説明するのに20主成分を要し、単純な線形圧縮では情報損失が大きいことが示された(図3)。

図2 次元削減手法の検証

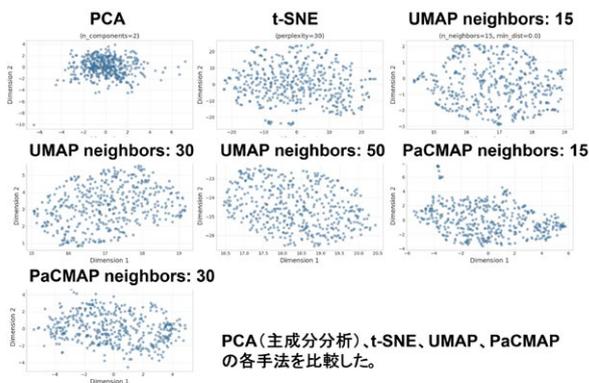
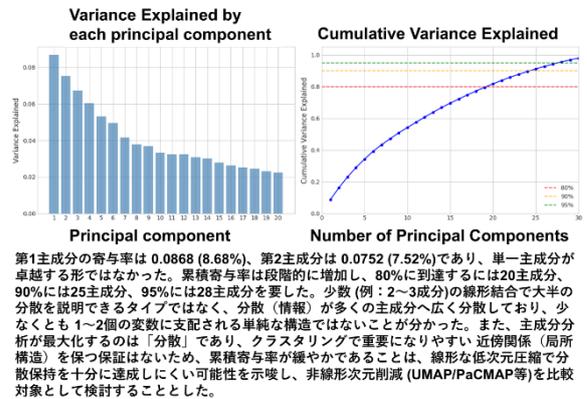


図3 PCAの累積寄与率



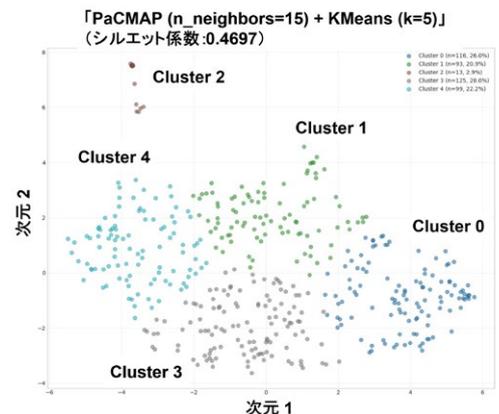
モデルの決定に関して、HDBSCANは内部指標が高いものの、ノイズ率が平均39.8%と高く、全例の層別化には不適であった。対照的に、PaCMAP(n_neighbors=15)とK-Means(k=5)の組み合わせは、シルエット係数(0.4697)、DBI(0.6292)、CHI(508.29)のすべてで最良のスコアを記録し、ノイズ率も0であったため、これを次元削減及びクラスタリング手法の最適組み合わせとして採択した。

4. 5つの臨床的フェノタイプの同定

最適モデルにより、以下の5つのクラスターが同定された(図4, 表1-2)。

- クラスター0 (N=116, 26.0%) : サルコペニア・アーチファクト群(低BMI・低アルブミン・高CRP、GNRI/PNI最低、eGFRは見かけ上高い)eGFRが79.45と見かけ上高いが、低BMI(19.10)、低ALB(2.92)、高CRP(3.58)を特徴とした。栄養指標であるGNRI(79.64:高リスク)およびPNI(35.48:重度)が最も

図4 クラスターの同定



低く、筋肉量低下によりクレアチニン値が低く抑えられている「偽りの腎機能良好」の可能性が示唆された。

- **クラスター 1 (N=93, 20.9%)** : メタボリック・リザーブ群 (BMI高値、脂質異常・糖尿病合併多しているが栄養状態は良好) BMIが22.51と最も高く、脂質異常症 (61%) や糖尿病 (43%) の合併が多かった。一方、栄養状態は良好 (GNRI 92.49) で、炎症 (CRP 2.47) も相対的に抑えられていた。
- **クラスター 2 (N=13, 2.9%)** : 構造的・社会的フレイル群 僧帽弁狭窄症や怠薬 (46%) が特異的に見られる少数群であった。
- **クラスター 3 (N=125, 28.0%)** : 心房性心筋症/AF-MR軸群 (女性90%、AF・MR高頻度) 女 (90%) が極めて多く、心房細動 (58%) や僧帽弁閉鎖不全症 (87%) を病態の中心としていた。
- **クラスター 4 (N=99, 22.2%)** : 心腎貧血/鬱血群 (重度腎不全・貧血・高度拡張障害・浮腫) 重度の腎不全 (eGFR 28.56)、貧血 (Hb 9.82)、高度な拡張障害 (E/e' 17.61) および浮腫 (59%) を呈する、古典的な心腎連関群であった。

表1 同定された5つのクラスターとその特徴量(1)

Variable	Overall	Cluster 0	Cluster 1	Cluster 2	Cluster 3	Cluster 4	p-value
n	Total	116	93	13	125	99	
Demographics							
Age, y	89.2 (6.1)	88.3 (6.6)	87.0 (6.4)	89.9 (4.6)	90.7 (5.6)	90.5 (5.3)	<0.01
Female, %	76%	68%	72%	85%	90%	70%	0.02
BMI, kg/m ²	20.4 (3.7)	19.1 (2.8)	22.5 (4.5)	18.1 (1.6)	19.8 (3.2)	21.2 (3.6)	<0.01
Echocardiography							
LVEF, %	60.4 (13.6)	59.7 (13.8)	63.0 (10.0)	59.9 (11.8)	56.6 (15.2)	63.6 (13.4)	<0.01
E/e'	14.2 (6.4)	12.6 (4.6)	13.0 (4.7)	14.5 (6.1)	13.7 (4.8)	17.6 (9.5)	<0.01
DcT, ms	215.5 (84.8)	213.3 (75.5)	217.3 (67.7)	279.9 (193.9)	195.6 (82.5)	233.1 (84.0)	<0.01
SVi, mL/m ²	33.5 (16.5)	30.7 (16.4)	31.0 (12.1)	38.2 (14.5)	31.3 (13.7)	41.6 (20.7)	<0.01
Comorbidities							
Atrial fibrillation, %	41%	39%	33%	38%	58%	29%	<0.01
Diabetes, %	20%	19%	43%	15%	7%	17%	<0.01
Hypertension, %	65%	48%	70%	69%	70%	72%	<0.01
Dyslipidemia, %	17%	4%	61%	23%	0%	13%	<0.01
Renal dysfunction, %	78%	25%	92%	100%	100%	98%	<0.01
Hyperuricemia, %	20%	1%	25%	8%	7%	56%	<0.01
AS, %	33%	40%	42%	62%	31%	49%	<0.01
AR, %	52%	55%	23%	46%	60%	66%	<0.01
MR, %	73%	73%	46%	77%	87%	78%	<0.01
MS, %	2%	0%	0%	62%	0%	0%	<0.01
Edema, %	39%	38%	27%	38%	32%	59%	<0.01
Non-compliance, %	2%	1%	0%	46%	0%	0%	<0.01

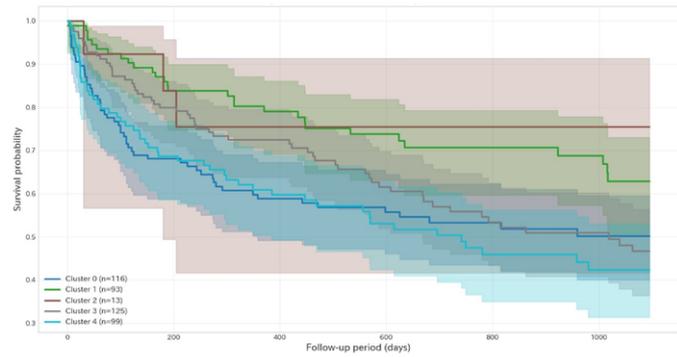
表2 同定された5つのクラスターとその特徴量(2)

Variable	Overall	Cluster 0	Cluster 1	Cluster 2	Cluster 3	Cluster 4	p-value
Laboratory Data							
eGFR, mL/min	48.2 (31.7)	79.5 (43.6)	41.3 (15.8)	35.7 (13.8)	41.2 (12.7)	28.6 (13.5)	<0.01
Na, mmol/L	138.5 (5.0)	137.3 (5.9)	138.0 (4.3)	139.8 (2.3)	139.4 (4.8)	139.4 (4.7)	0.01
K, mmol/L	4.1 (0.6)	3.9 (0.6)	4.3 (0.6)	4.0 (0.6)	4.0 (0.4)	4.4 (0.8)	<0.01
Hb, g/dL	10.9 (1.9)	11.1 (1.6)	11.3 (1.8)	10.9 (1.0)	11.2 (1.8)	9.8 (2.1)	<0.01
AST, U/L	53.9 (568.9)	25.1 (15.2)	162.2 (124)	23.5 (7.2)	24.1 (10.3)	27.6 (26.0)	0.86
ALT, U/L	28.5 (204.6)	17.6 (24.2)	73.1 (445.2)	14.0 (7.6)	14.5 (11.8)	18.9 (27.6)	0.80
CRP, mg/L	2.1 (3.2)	3.6 (4.2)	2.5 (3.7)	1.3 (1.8)	0.9 (1.3)	1.4 (2.0)	<0.01
ALB, g/dL	3.2 (0.5)	2.9 (0.5)	3.4 (0.5)	3.3 (0.4)	3.3 (0.4)	3.1 (0.6)	<0.01
Platelet, 10 ³ /μL	195.9 (76.5)	212.0 (89.1)	223.0 (90.5)	215.8 (58.5)	179.8 (48.2)	169.2 (63.8)	<0.01
BUN, mg/dL	28.6 (17.3)	20.1 (9.9)	27.7 (15.7)	25.6 (7.2)	25.9 (10.7)	43.4 (23.2)	<0.01
Cl, mmol/L	102.0 (5.7)	99.7 (6.2)	101.8 (5.2)	103.5 (3.2)	102.1 (4.7)	104.4 (6.0)	<0.01
TP, g/dL	6.5 (0.7)	6.3 (0.7)	6.7 (0.7)	6.6 (0.5)	6.6 (0.6)	6.5 (0.7)	<0.01
Lymphocyte, /μL	1261 (611)	1260 (669)	1449 (640)	1222 (351)	1280 (638)	1068 (431)	0.02
GGT, U/L	37.1 (50.4)	28.3 (24.1)	57.6 (91.1)	25.9 (15.5)	32.6 (28.2)	35.3 (38.8)	0.18
Cholesterol, mg/dL	163.9 (40.6)	155.5 (37.2)	165.0 (42.2)	170.7 (39.8)	168.3 (37.1)	166.1 (46.2)	0.18

5. 生存時間分析と解析系の安定性評価

3年間の追跡調査において、5つのクラスターの全死因死亡に関して、ログランク検定で解析したところ、有意差が認められた ($p=0.018$) (図5)。特に、クラスター1はクラスター0に対し、年齢・性別を調整した多変量Cox回帰分析においても有意に予後が良好であった ($HR=0.57$, $95\%CI: 0.445-1.029$, $p=0.014$)。また、解析系の安定性についてリサンプリング (Pairwise subsampling) によるクラスター割り当ての一致度は中程度 ($ARI=0.424$) であり、症例単位では境界領域での揺らぎが認められた。一方で、異なる手法 (Ward、GMM) 間での特徴量プロファイルは極めて高い相関 ($r>0.9$) を示しており、臨床的なフェノタイプ解釈の再現性は十分に担保されていることを確認した。

図5 クラスター別の生存時間分析

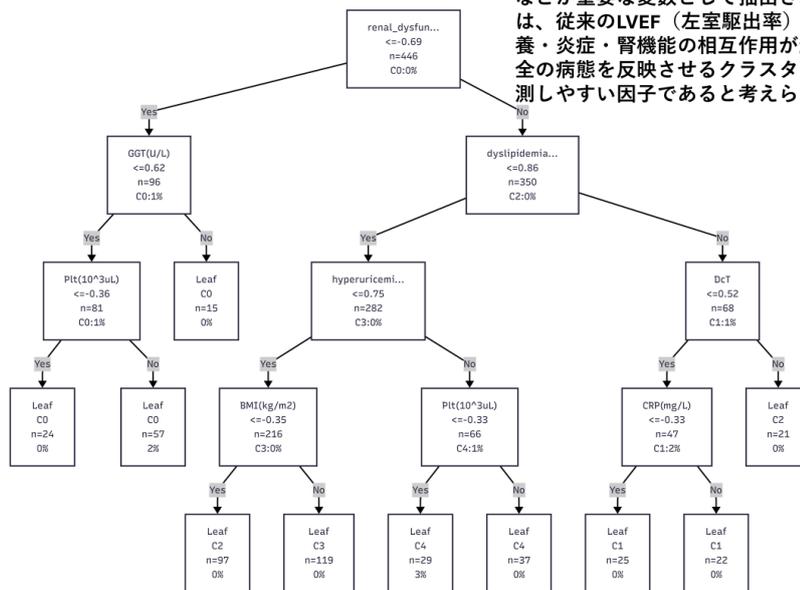


生存時間分析において、クラスター間で全死因死亡に有意差が認められた ($p=0.018$)。さらに、Cox回帰分析でクラスター0を基準とした場合、クラスター1は多変量解析においても有意に予後が良好であることが示された (ハザード比 0.57 , $p=0.014$)。

6. 事後解析：決定木とSHAPによる重要度

クラスター分類を目的変数とした代理モデルによる説明性解析 (決定木の可視化 (図6-7)、LightGBMモデルによるSHAP解析 (図8)) を用いた。eGFR <60 、BMI、CRP、脂質異常症、Hbなどがクラスター予測における重要変数として抽出され、4章で上述したフェノタイプの説明と整合する結果となった。

図6 決定木分析



クラスター分類の要因を可視化した結果、eGFR <60 、BMI、CRP、脂質異常症、Hbなどが重要な変数として抽出された。これは、従来のLVEF (左室駆出率) 以上に、栄養・炎症・腎機能の相互作用が超高齢心不全の病態を反映させるクラスター分類を予測しやすい因子であると考えられる。

図7 決定木分析：特徴量を決定する因子

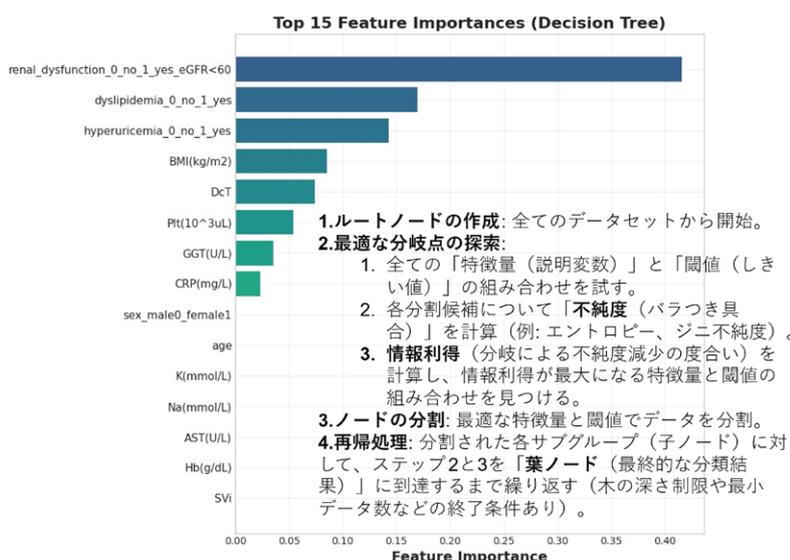
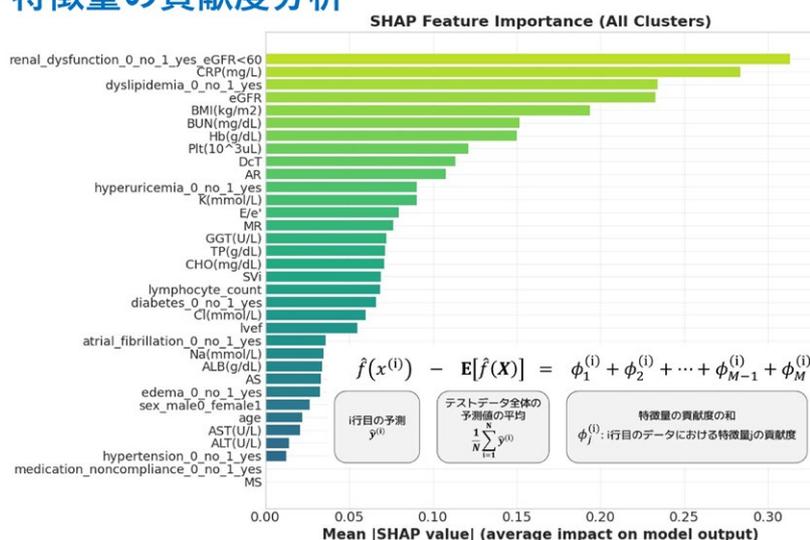


図8 SHapley Additive eXplanations (SHAP)による特徴量の貢献度分析



考察

本研究において臨床上意義の高い知見は、超高齢心不全患者において、従来のLVEF値以上に栄養・炎症・腎機能の三者の相互作用が表現型決定に寄与していることが判明した点である。特にクラスター0（サルコペニア・アーチファクト群）において、eGFRが高いにもかかわらず予後が不良であった事実は臨床上の新たな視点をもたらさう。これは、低栄養と炎症（高CRP）に伴う筋肉量減少がクレアチニン産生を低下させ、腎機能を過大評価させている可能性を強く示唆している。こうした「見かけ上の腎機能良好」に惑わされず、低BMIや低ALB、そしてGNRI等の栄養指標を併せて評価する必要があると考えられる。

一方で、クラスター1で見られた「メタボリック・リザーブ（肥満パラドクス）」の傾向は、超高齢者においては適度なエネルギー蓄積や栄養状態の維持が、心不全の異化代謝亢進に対

する保護因子として機能している可能性を支持していると考えられた。

手法面では、PCAの累積寄与率が低かったことは、臨床データが単純な線形関係では説明できない複雑な構造を持つことが判明した。今回、PaCMAPのような非線形手法を採用し、105通りの比較を経てモデルを選定したプロセスは、解析結果の客観性を担保する上で不可欠であった。本研究の安定性指標（ARI=0.424）の結果は、超高齢心不全の病態が不連続なクラスではなく、連続的なスペクトラム（連続体）であることを示唆している。しかし、手法を変えても特徴量の傾向は一貫して再現されたことから、本研究で同定したフェノタイプは、手法の恣意性に依存しない普遍的な病態構造を捉えていると考えられる。

要 約

本研究は、高齢心不全患者を対象に、最適な機械学習を用いて5つのフェノタイプを同定した。超高齢心不全では、従来のLVEF単独による分類より栄養・炎症・腎機能の相互作用を考慮し、総合的に表現型を把握することが重要であることが示された。特に、サルコペニア群では筋肉量減少によるクレアチニン低下で腎機能が過大評価されやすく、GNRIなどの栄養指標の併用が判別に重要であることが解った。メタボリック・リザーブ群に見られる「肥満パラドックス」は、適度なエネルギー蓄積が保護的に働く可能性を示唆していた。非線形次元削減の必要性と心不全病態の連続スペクトラム性も確認され、本研究の結果は、高齢心不全の個別化医療に向けた新たなフェノタイプングと治療指針の決定に貢献するものである。また、今回開発した解析方法は、解析ノイズを最小限に抑え、高齢者心不全集団を層別化するシステムであり、複雑な合併症を持つ超高齢者の個別化医療を現実的にする第一歩となる可能性がある。将来的には、これらのクラスターの臨床的意義を更に検証するために外的Validation研究が必要である。

文 献

1. Circulation 2014;63:e57-185.
2. American heart journal 2010;159:949-955.
3. JACC 2018;71:1921-1936.

謝 辞

本研究は、公益財団法人大和証券財団の第51回調査研究助成（課題名：高齢慢性心不全患者の大規模データベース構築と人工知能を用いた心不全病態の解明）により支援された。ここに厚く御礼申し上げます。