

開発テーマ

新型コロナウイルス対策のための重症化予測マーカー探索と患者の層別化

概要

患者の重症化回避のための個別化医療として、離散型特徴選択による重症化予測マーカー探索と離散型識別による患者の層別化を行う。これにより、データに基づく科学的根拠のある新型コロナウイルス対策を推進し、医療崩壊を防ぐ。

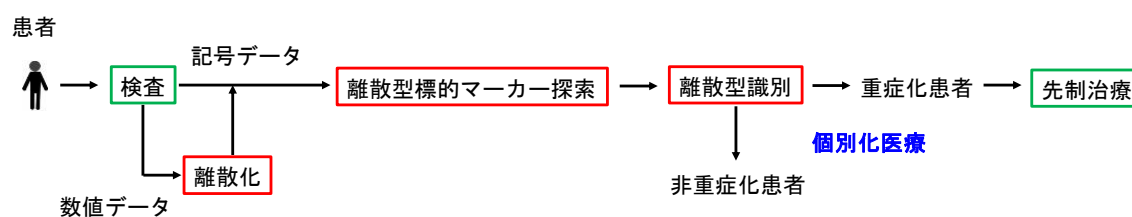
1. 研究開発の目的

新型コロナウイルス対策は、ICTの未整備に起因してデータの流通が滞り、現場医師の経験と知識に依存した人海戦術となって苦戦している。このため、データトランスフォーメーションを推進し、データに基づく科学的根拠のある対策が求められている。特に関心があるのは、多大な医療リソースを必要とする重症患者への対応である。医療崩壊の危機にあって、限られた医療リソースを枯渇させないためにも患者を重症化させないことが喫緊の課題である。そこで、個別化医療として、重症化する患者とそうではない患者を層別化（識別）して、重症化する患者に対しては重症化前に効果的な先制治療を施す。これは、重症化回避となって医療従事者の負担低減に直結し、医療の効率化において意義がある。

本研究開発の目的は、特徴選択である離散型標的マーカー探索により極めて多数の候補マーカーの中から有用な重症化予測マーカー群（マルチマーカー）を探索すること、更に重症化予測マーカー群を用いた離散型識別により高精度に患者の層別化を行うことである。

2. 研究開発の方法

本研究開発は、教師あり学習に基づき、まず重症化患者のクラスと重症化しない患者のクラスを構成し、次に両クラスを候補マーカー上で対比させて患者の層別に有用な重症化予測マーカー群を探索する。続いて重症化予測マーカー群を用いて患者の層別を行う。以下に処理の流れを示す。



重症化回避の流れ

2.1 本研究開発の独創性

一般に、医学データは量的データ（数値データ）と質的データ（記号データ）が混在し、このことがマーカー探索に制限を与えている。多くの解析手法では、数値データのみを対象としている。しかし、医学分野では性差、陽性/陰性、遺伝子変異の有/無等の名義尺度データと呼ばれる記号データは不可欠である。このため、機械的に名義尺度データを 0/1 に数値化することも考えられるが、本来記号としてしか意味のない名義尺度データの数値化には解釈上の問題がある。

本研究開発では、数値データを離散化し「医学的に意味のある記号データ」に変換する。このように全ての医学データを記号データに統一してから、数値データを対象とする従来の統計的パターン認識に対応する、独自の記号データを対象とした離散型標的マーカー探索と離散型識別を実現する。特筆すべきは、解析の結果である標的マーカー群を医師が解釈し、医学的な知見を獲得できる点にある。なお、数値データの離散化は、丸め処理に相当し、個人情報保護に役立つという長所を有する。また本手法には、数値データの解析と異なり、「次元の呪い」という高次元に起因する計算の困難さはない。

2.2 重症化予測マーカー群の探索

マーカー探索とは、統計的パターン認識では特徴選択とも呼ばれ、極めて多数の候補マーカーの中から予測・診断に有用な少数のマーカーを選択することである。病態が解明されていない新型コロナウイルスでは特に、候補マーカーに見落としかないように、意識して網羅的に、できるだけ数多くの候補を対象としなければならない。そのため、血液検査等の候補マーカーだけではなく、オミックス情報や遺伝子多型等の遺伝子情報を提供する候補マーカーも対象とする。

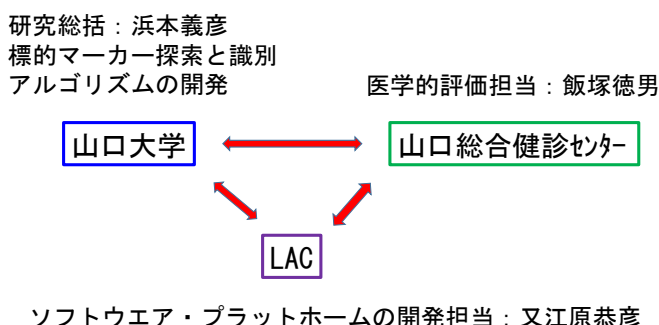
統計的パターン認識の特徴選択理論によれば、個々の標的マーカーではなく、標的マーカーの組合せが重要であることが知られている。このため、本手法では標的マーカー群を標的マーカーの組合せとして選択する。

2.3 患者層別化

重症化予測マーカー群を用いて、患者を重症化するか否かに識別する。記号データを対象とするため、独自の離散 Bayes 識別則を用いる[3] [4]。

離散 Bayes 識別則では、重症化予測マーカー群上で患者から得られる記号データを成分とする多次元ベクトルを入力し、各クラスの事後確率を計算して、事後確率が最大のクラスへ患者を識別する。

2.4 研究開発の体制



3. 研究開発の遂行能力

3.1 これまでの本研究開発に関連する研究

代表者は、これまで臨床医と共同で主に癌を対象とした個別化医療の実現を目指し、遺伝子発現データを用いた術後の肝癌の早期再発予測[1]から始まり、現在注目されているリキッドバイオプシ研究のさきがけとなる血液検査による腫瘍マーカーを用いた肝癌の高精度診断[2]がある。最近では離散 Bayes 識別則により、保険適用の臨床データを用いた術後の肝癌の早期再発予測[3]や早期胃癌のリンパ節転移予測[4]の成果がある。また予後予測として、抗癌剤の投与前の治療効果予測や術後管理で重要な術後合併症予測の研究にも取り組んでいる。

学術論文

- [1] N. Iizuka and Y. Hamamoto et al., Oligonucleotide microarray for prediction of early intrahepatic recurrence of hepatocellular carcinoma after curative resection, *Lancet*, Volume 361, Issue 9361, 15 Pages 923-929, March 2003.
[https://doi.org/10.1016/S0140-6736\(03\)12775-4](https://doi.org/10.1016/S0140-6736(03)12775-4)
- [2] N. Iizuka and Y. Hamamoto et al., Efficient detection of hepatocellular carcinoma by a hybrid blood test of epigenetic and classical protein markers, *Clinica, Chimica, Acta*, Vol.412, Issues1-2, pp.152-158, 2011.
<https://doi.org/10.1016/j.cca.2010.09.028>
- [3] H. Ogihara, N. Iizuka, Y. Hamamoto, Prediction of early recurrence of liver cancer by a novel discrete Bayes decision rule for personalized medicine, *BioMed Research International*, Vol.2016, Article number 8567479, 2016.
<https://doi.org/10.1155/2016/8567479>
- [4] A. Goto, J. Nishikawa, and Y. Hamamoto et al., Lymph node metastasis can be determined by just tumor depth and lymphovascular invasion in early gastric cancer patients after endoscopic submucosal dissection, *European Journal of Gastroenterology and Hepatology*, Vol.29, Issue 12, pp.1346-1350, 2017.
<https://doi.org/10.1097/MEG.0000000000000987>

特許

特許第 6041331 号 (PCT 出願の JST 支援 : PCT/JP2016/088734)

3.2 本研究開発に関連する科研費（代表のみ）の獲得実績

①2015-2017 基盤研究 (C)

「離散 Bayes 識別理論による個別化医療のための肝癌再発予測」

②2012-2014 基盤研究 (C)

「ブール代数的パターン認識によるゲノムデータと臨床データを統合した肝癌再発予測」

③2009-2011 基盤研究 (C)

「少数サンプルを対象とした統計的パターン認識とその肝癌診断への応用」

④2006-2008 基盤研究 (C)

「標的蛋白質を用いたパターン認識による癌診断システムの開発」

⑤2004-2005 基盤研究 (C)

「マイクロアレイを用いた統計的パターン認識による癌の転移予測システムの開発」

3.3 未来技術社会実装事業（令和 2 年度採択）

レジリエントで持続可能な社会を創る「スマートシティ宇部プロジェクト」

事業：AI 活用型個人向けヘルスケアサービスの提供

体制：山口大学、山口総合健診センター、LAC

健診データ等のプラットフォームを構築し、アプリによる情報提供や、AI 解析システムによる健康状態予測サービスをプッシュ通知により提供

4. 学術の現状

現在の AI は、深層学習とも言われ、画像診断に特化している。しかし、数値データと記号データが混在した、血液検査、遺伝子多型などの一般的なマーカーに対する標的マーカー探索や患者層別化のアルゴリズムの研究は見当たらない。

5. 本研究開発の可能性

新型コロナウイルスにおけるリスクの態様に応じてメリハリの効いた対策を適切に講じることによって、重症者をできる限り抑制しつつ、社会経済活動を継続することが可能になる。このために活用される本手法は、今後想定される新興ウイルスに関しても対処可能な汎用性の高い手法である。

5.1 二次感染を起こしやすい患者の同定

コロナ感染者のほとんどは他者に感染させることがなく、ごく一部の感染者が多数の二次感染者を発生させることが知られている。医療の効率化のためにも、このような二次感染を起こしやすい人を迅速に同定することが必要となる。そこで、二次感染を起こしやすい、感染リスクの高い患者のクラスと感染リスクの低い患者のクラスを構成し、有用な標的マーカー群を探索して感染リスクの高い患者を同定する。この他に、退院後の予後不良の患者の同定も可能である。

本手法によれば、重症化対非重症化、感染リスク大対感染リスク小、予後良好対予後不良等のような、様々な相対する 2 クラスに対する標的マーカー群探索と患者の層別化が可能である。

5.2 標的マーカー群の検討によるメカニズムの解明

重症化や感染リスクの有無等、クラス間の相違を顕著に表す標的マーカー群自体に医学的な価値がある。探索された標的マーカー群を調べることにより、重症化や感染のメカニズムを解明する手がかりが得られる。

5.3 仮説生成としての AI の活用

AI は、世間では万能のごとく誤解されているが、本研究開発ではヒトの臨床推論力を高める一つのツールという位置づけである。例えば、解析に用いたデータ上で高精度に予測できる標的マーカー群が求められたとすれば、これは AI を活用してヒトが重症化、感染モデルの仮説を発見したことを意味する。この仮説を更に多くのデータ上でシミュレーションにより検証することで、仮説の正当性は高められ、恣意性のない、データに基づく客観的な成果が得られる。