

「遺伝的リスクが低くても予防効果が高い」ケースも！

**ポリジェニックリスクスコア×機械学習で紐解く**

**生活習慣病の遺伝的リスクと予防効果との関係**

【研究成果のポイント】

- ◆ 機械学習<sup>\*1</sup>とポリジェニックリスクスコア<sup>\*2</sup>を用いて、生活習慣病のリスク因子の改善による疾患予防効果と遺伝的リスクとの関係性を評価した。
- ◆ 「冠動脈疾患の遺伝的リスク」と「喫煙の改善による予防効果」、および「2型糖尿病の遺伝的リスク」と「肥満の改善による予防効果」はそれぞれ正の相関を認め、これらの疾患とリスク因子の組み合わせにおいて、遺伝的高リスク集団で高い予防効果が期待されることが示唆された。
- ◆ 一方、他の生活習慣病とリスク因子の関係においては、必ずしもそのような高い正の相関を認めず、遺伝的リスクが低くても疾患の予防効果が高い人たちが多く存在しうることが示唆された。

❖ 概要

大阪大学大学院医学系研究科の内藤龍彦 助教(研究当時/現:マウントサイナイ医科大学/ニューヨークゲノムセンター 博士研究員)、岡田随象 教授(遺伝統計学/理化学研究所生命医科学研究センター システム遺伝学チーム チームリーダー/東京大学大学院医学系研究科 遺伝情報学 教授)、京都大学白眉センターの井上浩輔 特定准教授(社会疫学)らの研究グループは、**機械学習**と**ポリジェニックリスクスコア**を用いて、冠動脈疾患、2型糖尿病、脂質異常症、高血圧症の4つの生活習慣病とその主要なリスク因子である喫煙や肥満との関係が、その疾患のポリジェニックリスクスコアによってどのように変化するかを評価しました。その結果、「**冠動脈疾患の遺伝的リスク**」と「**喫煙の改善による疾患予防効果**」、および「**2型糖尿病の遺伝的リスク**」と「**肥満の改善による疾患予防効果**」にそれぞれ**高い正の相関がある**ことが解明されました。一方、**他の生活習慣病とリスク因子の関係においては、遺伝的リスクが高い人たちが必ずしもリスク因子の改善による予防効果が高いわけではなく、遺伝的リスクが低くてもリスク因子の改善による高い予防効果が期待される人たちが多く存在する可能性がある**ことも示されました。

本研究成果は、英国科学誌「Communications Medicine」(オンライン)に、9月20日(金)(日本時間)に公開されました。

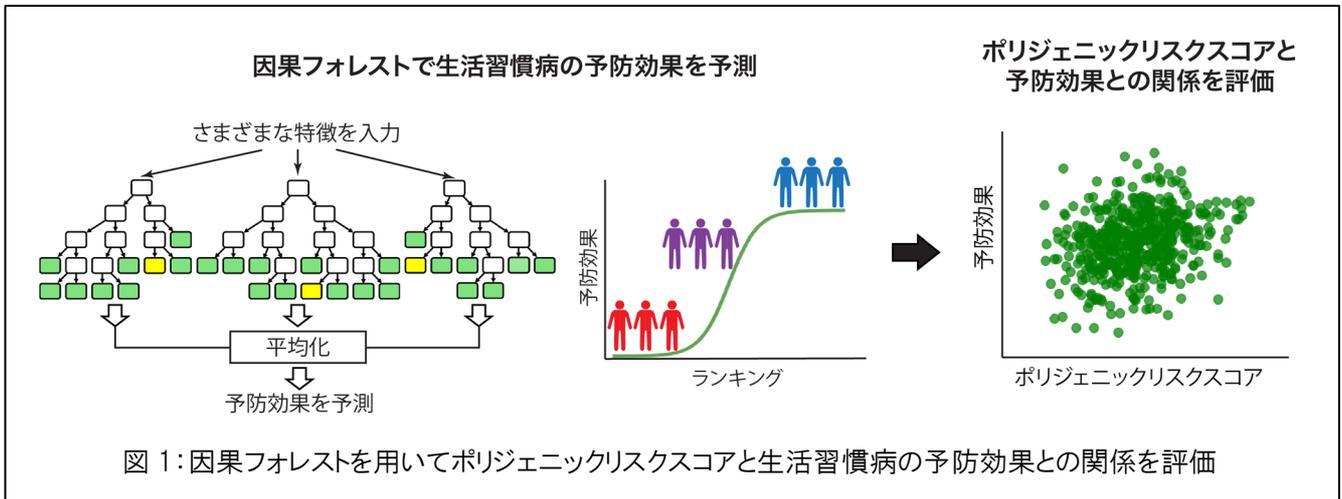
❖ 研究の背景

ゲノムワイド関連解析<sup>\*3</sup>により、心血管疾患や糖尿病など様々な生活習慣病の遺伝的因子が明らかになってきました。そして、ゲノムワイド関連解析の結果を活用することで、遺伝的因子による疾患発症リスクを単一のスコアであるポリジェニックリスクスコアとして計算できるようになりました。ポリジェニックリスクスコアは遺伝的にその疾患にかかりやすい人たち(遺伝的高リスク群)を予測できることから、個別化医療への将来的な応用が期待されています。

一方、近年の疫学研究では、ある疾患の高リスク群が、その疾患のリスク因子を改善することで期待される治療・予防効果が高い集団(高ベネフィット群)と必ずしも一致しない(=ある疾患の高リスク群において、治療や予防医療がどの程度効果的なのかがわからない)ことがわかってきました。そのため、リスクのみではなくベネフィットにも着目することの重要性が指摘されています。そして治療・予防効果がどのように異なるのか(効果の異質性)を把握することが、そのようなベネフィットに着目した医療を実現する上で直接的に役に立つ情報になります。これまで、生活習慣病のポリジェニックリスクスコアが高い集団(遺伝的高リスク群)が、喫煙や肥満といった生活習慣リスク因子を改善した際に高い予防効果が得られるかどうかは明らかではありませんでした。

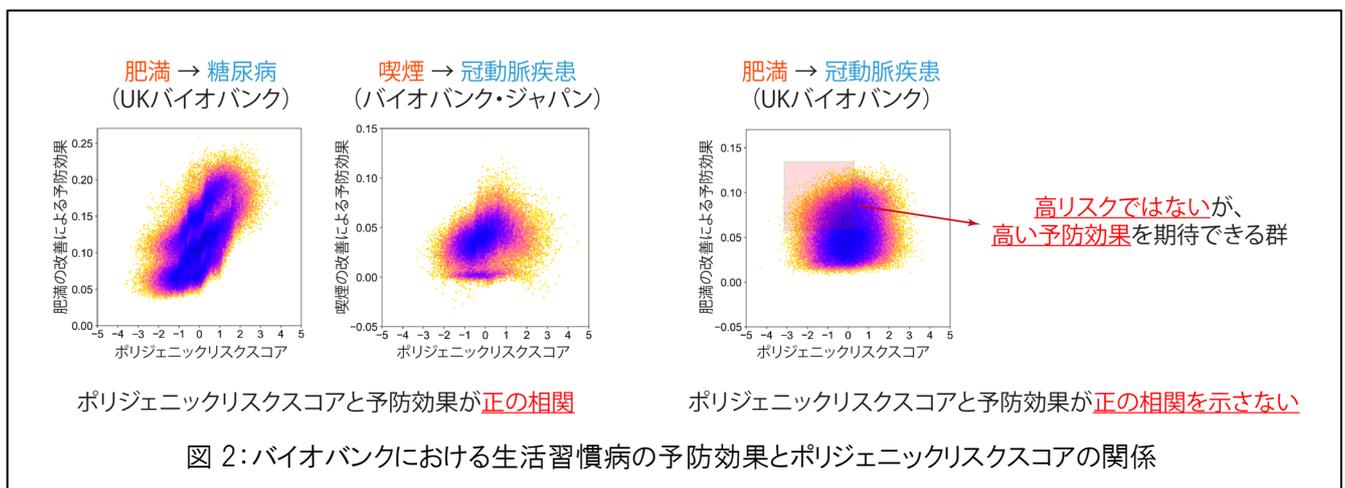
## ❖ 本研究の成果

研究グループは、機械学習法である因果フォレスト<sup>※4</sup>を用いて、生活習慣病の予防効果がポリジェニックリスクスコアによってどのように変化するかを評価しました(図 1)。因果フォレストは、ある介入の効果(観察研究においては曝露とアウトカム関連)を個人レベルで予測し、効果の異質性すなわち効果の個人間のばらつきを評価することができる手法です。



研究グループは、上記手法をバイオバンク・ジャパン<sup>※5</sup>と UK バイオバンク<sup>※6</sup>(英国)が保有するゲノム・臨床情報に適用することで、冠動脈疾患、2 型糖尿病、脂質異常症、高血圧症の 4 つの疾患とその主要なリスク因子である喫煙や肥満との関係が、その疾患のポリジェニックリスクスコアによってどのように変化するかを評価しました。結果、バイオバンク・ジャパンでは「冠動脈のポリジェニックリスクスコア」と「喫煙が冠動脈に与える影響」に、UK バイオバンクでは「2 型糖尿病のポリジェニックリスクスコア」と「肥満が 2 型糖尿病に与える影響」に強い正の相関がみられました(図 2)。この結果は、これらの疾患と生活習慣リスク因子の組み合わせにおいて、ポリジェニックリスクスコアが高い群では、生活習慣リスク因子を改善した際に高い疾患予防効果が見込める可能性を示唆しています。

一方、その他の疾患とリスク因子の組み合わせにおいては、ポリジェニックリスクスコアとリスク因子による疾患予防効果との正の相関が必ずしもみられません(図 2)。この結果は、ポリジェニックリスクスコアが高い人たちにおいて、それらのリスク因子を改善した場合でも、必ずしも高い疾患予防効果が得られるとは限らないことを示唆しています。言い換えると、ポリジェニックリスクスコアが低くても、生活習慣の改善によって高い疾患予防効果が期待できる人たちが多く存在すると解釈することができます。



## ❖ 本研究成果が社会に与える影響(本研究成果の意義)

本研究では、機械学習を用いて生活習慣病の遺伝的リスクと予防効果との関係性を評価し、生活習慣病の遺伝的リスクと予防効果との関係が疾患やリスク因子により様々なパターンを呈示することが示されました。機械学習を活用した将来的なゲノム個別化医療の礎になることが期待されます。

## ❖ 特記事項

本研究成果は、2024年9月20日(木)(日本時間)に英国科学誌「Communications Medicine」(オンライン)に掲載されました。

【タイトル】 “Machine learning reveals heterogeneous associations between environmental factors and cardiometabolic diseases across polygenic risk scores.”

【著者名】Tatsuhiko Naito<sup>1-3,12,13</sup>, Kosuke Inoue<sup>4,5,12</sup>, Shinichi Namba<sup>1,3,6</sup>, Kyuto Sonehara<sup>3,6</sup>, Ken Suzuki<sup>1</sup>, BioBank Japan<sup>※</sup>, Koichi Matsuda<sup>7</sup>, Naoki Kondo<sup>4</sup>, Tatsushi Toda<sup>2</sup>, Toshimasa Yamauchi<sup>8</sup>, Takashi Kadowaki<sup>9</sup>, Yukinori Okada<sup>1,3,6,10,11-13</sup>

### 【所属】

1. 大阪大学 大学院医学系研究科 遺伝統計学
2. 東京大学 大学院医学系研究科 神経内科学
3. 理化学研究所 生命医科学研究センター システム遺伝学チーム
4. 京都大学 大学院医学系研究科 社会疫学
5. 京都大学 白眉センター
6. 東京大学 大学院医学系研究科 遺伝情報学
7. 東京大学 大学院新領域創成科学研究科 メディカル情報生命専攻 クリニカルシーケンス分野
8. 東京大学 大学院医学系研究科 糖尿病代謝内科
9. 虎の門病院
10. 大阪大学 免疫学フロンティア研究センター(IFReC) 免疫統計学
11. 大阪大学ヒューマン・メタバース疾患研究拠点(PRIMe)
12. 共同筆頭著者
13. 共同責任著者

※ 各バイオバンク研究グループ全員のリストは論文中に記載

DOI: <https://doi.org/10.1038/s43856-024-00596-7>

本研究は、JSPS 科研費 JP 22K17392, 22H00476、国立研究開発法人日本医療研究開発機構(AMED) JP22rea522107, JP22ek0410075, JP23km0405211, JP23km0405217, JP23ek0109594, JP23ek0410113, JP223fa627002, JP223fa627010, JP233fa627011, JP23zf0127008, JP24tm0424228、国立研究開発法人科学技術振興機構(JST)ムーンショット型研究開発事業 JPMJMS2021, JPMJMS2024、武田科学振興財団、大阪大学大学院医学系研究科バイオインフォマティクスイニシアティブ、大阪大学感染症総合教育研究拠点(CiDER)、大阪大学ワクチン開発拠点先端モダリティ・DDS 研究センター(CAMaD)、L-INSIGHT、日本応用酵素協会の支援を受けて行われました。

## ❖ 用語説明

### ※1 機械学習

人工知能の一種で、大量のデータを学習することで、分類・予測や特徴抽出を行うコンピューターアルゴリズム。

### ※2 ポリジェニックリスクスコア

各個人の持つ遺伝的なリスクをスコア化して、疾患の発症や進展を予測する手法。GWAS の結果から算出される各遺伝的変異のその疾患に対する効果量を、各個人の持つ遺伝子型を重み付けして足し合わせることで計算される。

### ※3 ゲノムワイド関連解析

Genome-wide association study: GWAS。

ヒトゲノム配列上に存在する数百万～数千万か所の遺伝子多型と疾患の発症の関係を網羅的に検定することで、疾患の発症に関わる遺伝子多型を特定する遺伝統計解析手法。

### ※4 因果フォレスト

機械学習モデルの一種であるランダムフォレストを用いた因果推論手法の一つ。個人が有する(年齢・性別などの)特徴に基づいて因果の効果を予測し、効果の異質性を評価することができる。

### ※5 バイオバンク・ジャパン

日本人集団約 27 万人を対象とした生体試料バイオバンクであり、ゲノム DNA や血清サンプルを臨床情報と共に収集し、研究者へのデータの公開や分譲を行っている。

### ※6 UK バイオバンク

英国の約 50 万人を対象とした国家的な生体試料バイオバンク。ゲノム情報や多彩な臨床情報、追跡情報を収集し、世界中の研究者にデータの公開や分譲を行っている。

#### 【研究者のコメント】<内藤龍彦 助教(研究当時)>

ポリジェニックリスクスコアに関する研究は、予測精度の評価や手法開発に焦点を当てたものが多く、スコアの高い人に対して治療や予防医療を集中的に行うべきかどうかという問題については多くは検討されていない印象でした。本研究が、そのような課題に対して新しい視点を提供し、遺伝情報を活用した将来的なゲノム個別化医療の実現に貢献できれば幸いです。本研究はバイオバンク・ジャパンと UK バイオバンクの保有するゲノム・臨床情報を活用することで実現できました。両バイオバンクの関係者の皆さまをはじめ、すべての共同研究者、研究支援機構、そして検体をご提供いただいた方々に深く感謝を申し上げます。