

2023年 5月 18日

報道機関 各位

【配信先】新潟県政記者クラブ、文部科学記者会、科学記者会

新潟大学

## 機械学習を用いた 神経難病の発症年齢予測を実施

新潟大学脳研究所脳神経内科学分野の畠野雄也非常勤講師、同研究所脳神経疾患先端治療研究部門の石原智彦特任准教授、同研究所脳神経内科学分野の小野寺理教授らの研究グループは、当同研究所脳神経内科で遺伝子検査が行われた脊髄小脳変性症3型（SCA3）（注1）292例と歯状核赤核・淡蒼球ルイ体萎縮症（DRPLA）（注2）203例のデータを用いて、キャリア（注3）の方の現在の年齢と遺伝情報（CAGリピート数）から、それぞれの年齢における未発症確率を予測しました。

また、キャリアの方の現在の年齢とCAGリピート数を入力することで、特定の年齢における未発症確率を図示できるWindows 64bit用アプリケーションをリリースしました。

本研究成果は、2023年5月4日、科学誌「Neurology Genetics」のオンライン版に掲載されました。

### 【本研究成果のポイント】

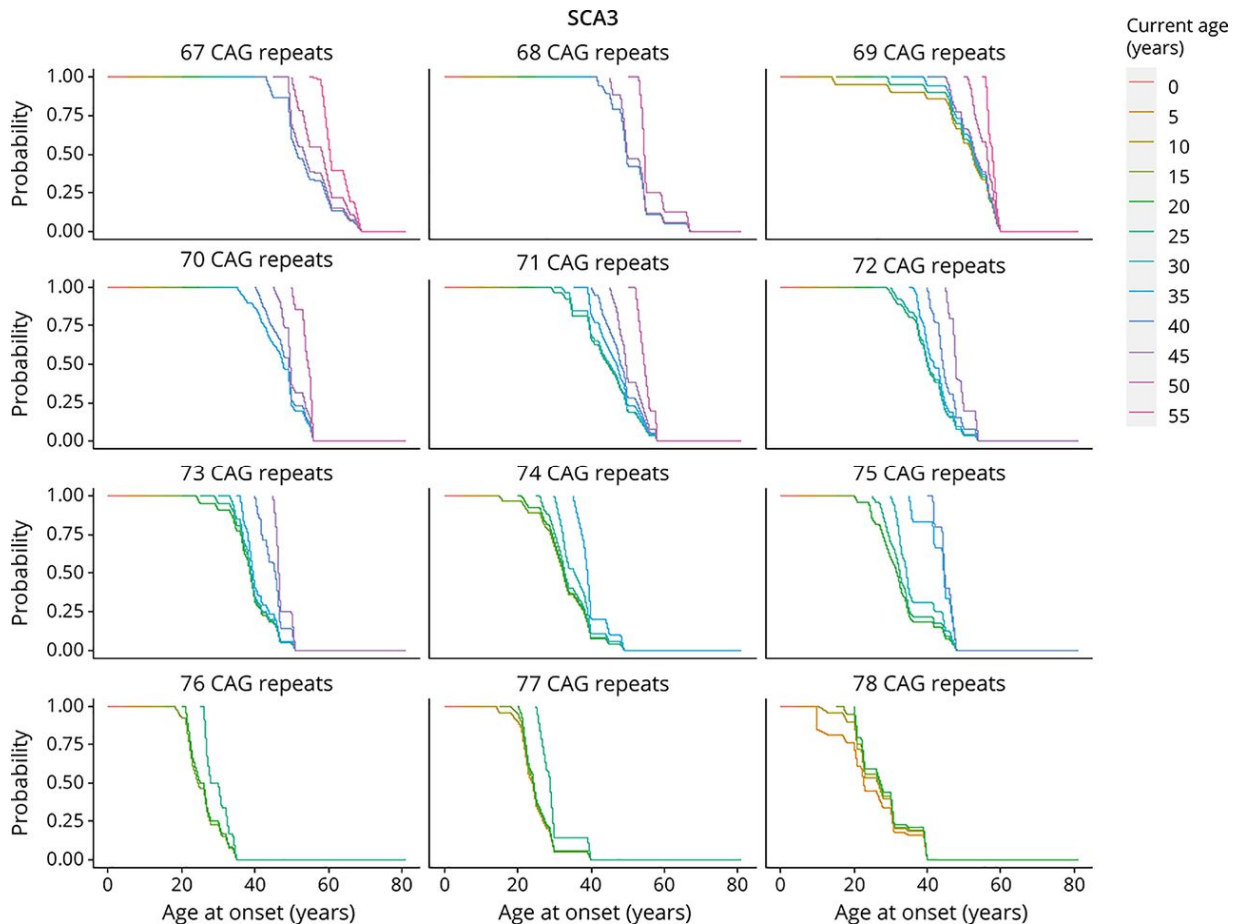
- 機械学習の生存分析を用いることにより、従来のパラメトリック生存分析（注4）よりもより正確に特定の年齢での未発症確率を予測しました。
- 未発症確率を図示するアプリは、Web上で公開しています  
(<https://github.com/yuya-hatano/SCA-onset>)。

### 1. 研究の背景

SCA3、DRPLAを含むポリグルタミン病は、原因遺伝子内のCAGリピートの異常伸長によっておこる病気です。ポリグルタミン病では、発症年齢とCAGリピート数が逆相関することが知られています。従来はパラメトリック生存分析を利用した発症年齢を予測が行われてきましたが、より精度の高い予測法が求められていました。我々は、2種類の機械学習の生存分析を使って発症年齢を予測し、予測精度を6種類のパラメトリック生存分析と比較しました。

## II. 研究の概要・成果

2種類の機械学習法（Random Survival ForestとDeepSurv）は、Integrated brier scoreという指標で、パラメトリック生存分析よりも良好に予測できました。特に、予測精度が高かったRandom Survival Forestで最終的な予測を行いました。



画像は Hatano Y, et al. Neurol Genet 2023 からの引用

## III. 今後の展開

本研究は、臨床現場や治験における活用が期待されます。今後は多施設で更に症例数を増やした解析を行い、より精度の高い発症確率の予測を目指します。

## IV. 研究成果の公表

本研究成果は、2023年5月4日、科学誌「Neurology Genetics」のオンライン版に掲載されました。

論文タイトル：Machine Learning Approach for the Prediction of Age-Specific Probability of SCA3 and DRPLA by Survival Curve Analysis

著者：Yuya Hatano, Tomohiko Ishihara, Sachiko Hirokawa, Osamu Onodera

doi: [10.1212/NXG.000000000200075](https://doi.org/10.1212/NXG.000000000200075)

### 【用語解説】

（注1）脊髄小脳変性症3型（SCA3）：顕性遺伝性の脊髄小脳変性症の一つです。日本に比較的多いタイプの脊髄小脳変性症です。

（注2）歯状核赤核・淡蒼球ルイ体萎縮症（DRPLA）：顕性遺伝性の脊髄小脳変性症の一つです。発症年齢によって症状が異なる特徴があります。

（注3）キャリア：ある病気の原因となる遺伝子の変異をもっているものの、その病気の症状が現れていない人のこと。

（注4）パラメトリック生存分析：パラメトリック手法は検定対象の母集団が何らかの分布（正規分布など）に従う前提で使用する統計学的な解析手法です。生存分析は何らかのイベントが起きるまでの時間とその発症率を調べる統計的手法のことです。

#### 本件に関するお問い合わせ先

新潟大学脳研究所脳神経疾患先端治療研究部門

特任准教授 石原智彦（いしはらともひこ）

TEL：025-227-0663

E-mail：ishihara@bri.niigata-u.ac.jp