

Reference-based standardization approach stabilizing small batch risk prediction via polygenic score

基準となる集団を利用してポリジェニックスコアを標準化し

小集団でも安定したリスク推定を可能にするアプローチ

Yoichi Sutoh,^{1,2} Tsuyoshi Hachiya,^{1,2} Yayoi Otsuka-Yamasaki,^{1,2} Tomoharu Tokutomi,^{3,4,5} Akiko Yoshida,^{3,4}
Yuka Kotozaki,^{6,7} Shohei Komaki,^{1,2} Shiori Minabe,^{1,2} Hideki Ohmomo,^{1,2} Kozo Tanno,^{6,7}
Akimune Fukushima,^{3,4,8} Makoto Sasaki^{9,10} and Atsushi Shimizu^{1,2*}

須藤洋一,^{1,2} 八谷剛史,^{1,2} 山崎弥生,^{1,2} 徳富智明,^{3,4,5} 吉田明子,^{3,4} 事崎由佳,^{6,7} 小巻翔平,^{1,2}

美辺詩織,^{1,2} 大桃秀樹,^{1,2} 丹野高三,^{6,7} 福島明宗,^{3,4,8} 佐々木真理,^{9,10} 清水厚志^{1,2*}

- 1 岩手医科大学 災害復興事業本部 いわて東北メディカル・メガバンク機構
生体情報解析部門
- 2 岩手医科大学 医歯薬総合研究所 生体情報解析部門
- 3 岩手医科大学 災害復興事業本部 いわて東北メディカル・メガバンク機構 イノベー
ション推進・人材育成部門
- 4 岩手医科大学 医学部 臨床遺伝学科
- 5 川崎医科大学 小児科学
- 6 岩手医科大学 災害復興事業本部 いわて東北メディカル・メガバンク機構 臨床研
究・疫学研究部門
- 7 岩手医科大学 医学部 衛生学公衆衛生学講座
- 8 北上済生会病院
- 9 岩手医科大学 災害復興事業本部 いわて東北メディカル・メガバンク機構
- 10 岩手医科大学 医歯薬総合研究所 超高磁場 MRI 診断・病態研究部門

*責任著者

【研究のポイント】

岩手医科大学 いわて東北メディカル・メガバンク機構 生体情報解析部門の須藤洋一 特命准教授、部門長の清水厚志 教授らの研究グループは、ポリジェニックスコア(PGS)^{*1} による病気リスクの推定を少人数でも信頼できる形で行うための手法を開発しました。

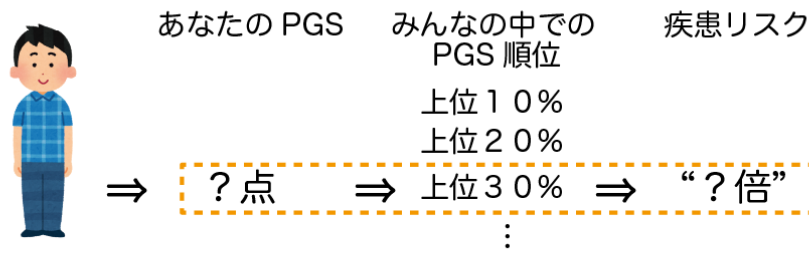
PGS による発症リスクの表し方は、一般的に、「集団の中で自分の PGS がどのくらいの高さにあり、その高さがあると、標準的な PGS を持つ方々と比べて何倍病気になりやすいのか」という形で表します (図1)。つまり、集団の PGS と自分の PGS をきちんと比べないと正しいリスクが算出できませんが、このとき、どのようなことに注意して比べるべきかなど、細かい部分が十分に検証されていませんでした。

研究グループが調べたところ、信頼性の高いリスク計算に大切なものは、1) 一般集団から、できるだけ多人数の PGS を得て比較すること、2) 一般集団の PGS と自分の PGS が同じ多型^{*2}を利用して計算されていること、の少なくとも2点あることがわかりました。

PGS は病気予防の目的で利用が期待されており、当機構も含め、世界中で検証が進められています。PGS に基づいた病気のリスクを受け取った方が、生活習慣の改善など、自身の病気予防に役立ててもらうためには、そもそも PGS が信頼性の高い形で計算されていることが前提になります。今回の結果は、その信頼性を確保するために必要な、2つの要点を明確に示したものになります。

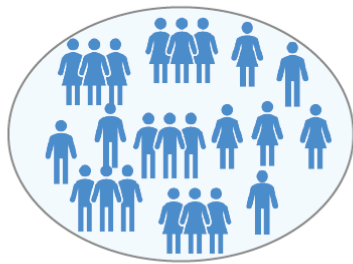
この成果は 2025 年 1 月 30 日づけで国際遺伝疫学協会 (International Genetic Epidemiology Society, IGES)の専門誌 *Genetic Epidemiology* に掲載されました。

PGS によるリスク推定では「集団中のどれくらいの位置にいるか」に基づいて疾患のリスクを推定する



安定したリスク推定結果が得られるために必要な条件を調べた

1. 順位を決めるとき、比べる相手の集団は大きい方が良い
2. 同じ多型を使って PGS を算出する必要がある



大きい集団の中での順位はぶれにくく、結果が変わりにくい



PGS を計算する際、個人と集団で同じ多型に基づいて計算するように揃える必要がある

図1 研究成果の概要

PGS を使った疾患リスクの推定を行う際に必要な条件について検証した。

【概要】

脳卒中、心疾患、糖尿病などの身近な疾患のかかりやすさは、運動習慣や食事習慣の改善により、低減させることができます。しかし、それとは別に、これらの疾患のかかりやすさには生まれ持った個人差があることもわかっています。こうした生まれ持った個人差は遺伝的な影響が大きいいため、これまでも病院などで家族の病歴をお聞きすることにより経験的に推定してきました。

近年、こうした生まれ持った病気のかかりやすさを客観的に数値化する方法として、ポリジェニックスコア(PGS)^{*1}という方法が開発され、盛んに研究が進められています。PGSは、人の遺伝情報の中で、ある病気に関連性のある多型^{*2}を予め特定しておき、各個人がそ

の多型を何箇所持っているかを調べ、得点化します。

PGS から、ある人の病気のリスクを推定する際は、この得られた PGS が、集団全体の中でどのくらいの順位にあるかをまず調べます (図 1)。集団全体の中で、例えば、PGS が上位 10%にある人はどのくらいの病気のリスクがあるのか、などは過去の研究ですでに調べてあるため、PGS の順位さえわかれば、病気のリスクを推定することができます。

発症前に自分自身の病気のかかりやすさがわかれば、より生活習慣に気をつけるなどの対処をすることができます。このため、PGS は近い将来の予防医学で用いられる指標として大変注目されており、現在も世界中で研究が進められています。

ただし、PGS を知って生活習慣を改めようと思うためには、PGS の値が科学的に十分信頼できるものである必要があります。例えば、測るたびに毎回変わってしまうような数値では、それを信じて生活を改めようとはなかなか思えないかもしれません。私達は、将来の予防医学のため、信頼できる PGS 計算に必要な要素を検証し、明確にしておく必要があると考えました。

まず、先ほど述べました通り、PGS による病気のリスク推定では、個人の PGS を調べた後に集団全体の中で自身の PGS がどの順位にあるかを知ることからリスクを推定します。では、この比べる集団はどのようなものが適切なのでしょうか。わたしたちは、この集団が大きければ大きいほど、推定の正確さが高まることをデータに基づいて明らかにしました。今回は集団の例として、岩手県・宮城県の地域住民の皆様にご協力いただいている東北メディカル・メガバンク計画地域住民コホートのデータ (最大 5 万 3 千人) 等を利用しましたが、こうした大規模な地域住民コホートのデータは、まさに比較する集団として適していると言えます。

比較する集団の大きさと その集団中での個人の PGS の順位推定のぶれ幅

比較する集団が大きいほどぶれ幅は小さくなる

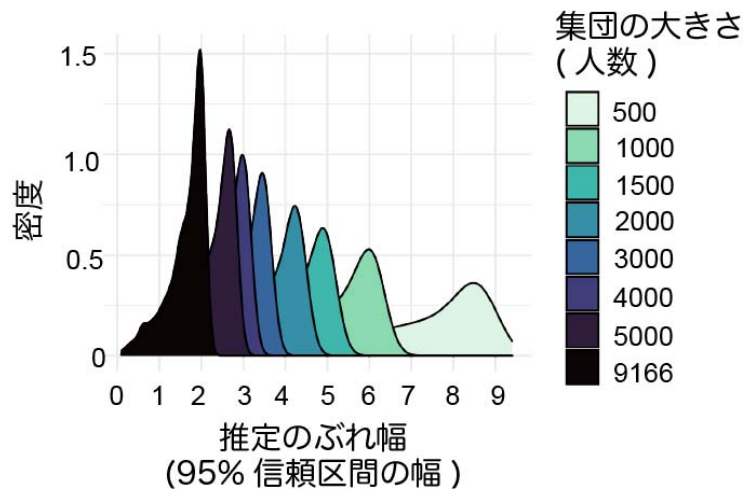


図2 比較する集団の大きさと推定のぶれ幅の関係

ある個人の PGS の順位を推定する際、大きい集団を基準に推定すると、推定のぶれ幅（誤差）が小さくなる。この図では、500 人の集団（薄緑）よりも、5000 人の集団（黒紫）と比較した方が、個人の PGS 推定順位のぶれ幅が小さい方に収まるようになる。

また、わたしたちは、個人の PGS と集団の PGS を比較するとき、多型の場所を揃えて PGS が計算してあることが重要であることも明らかにしました。PGS を計算する際に多くの多型を調べる必要がありますが、この際に使った実験方法や機材の種類によって、ごく一部の多型が計算から抜け落ちることがあります。この欠損データをそのままにして集団と個人の PGS を比較してしまうと、当然個人の PGS の順位を高く見積もったり、低く見積もったりするなど問題が生じます。また、計算から抜け落ちる部分は集団の性質によっても異なるため、対象の個人と比べる集団の性質があまりにかけ離れていると、それだけで推定結果が影響を受けてしまいます。今回の研究では、こうした抜け落ちる部分を事前に考慮し、個人と集団とできっちりと計算に使う多型を揃えてから PGS を計算し、比較を行うことで、機材や集団の性質に由来する影響をできるだけ小さくできることを明らかにしました。

【まとめと展望】

PGS によるリスク推定では、集団中の順位を正確に推定することが大切です。また、比べる集団が異なれば順位も異なる可能性があるため、「どの集団と比べたときの順位か」が明確にされていることも大切です。できるだけ大きな集団の中で自分の PGS の順位を推定することで、信頼性の高いリスク推定ができることを本研究は示しています。生活習慣改善のモチベーションを長く保っていただくためにも、推定の信頼性は重要な意味を持つと思われまます。

【用語解説】

*1 ポリジェニックスコア (PGS)

ポリジェニックリスクスコア (PRS; Polygenic Risk Score) とも呼ばれます。ある疾患のリスクと多型*2との関係をあらかじめ調べておき、その関係性の強さに応じて、各多型に得点を割り振っておきます。ある人の遺伝情報 (ゲノム) を調べた際、疾患と関係のある多型が見つかるごとに割り振られた得点を加算し、最終的な個人ごとの得点を出します。この得点が PGS です。計算に用いる多型は一人当たり数百万箇所にのぼることもあり、多くの場合、スーパーコンピュータ等の大型計算機が必要になります。

*2 多型

ヒトの遺伝情報 (ゲノム) の中には個人ごとに違いが見られる部分があり、これを多型と呼びます。