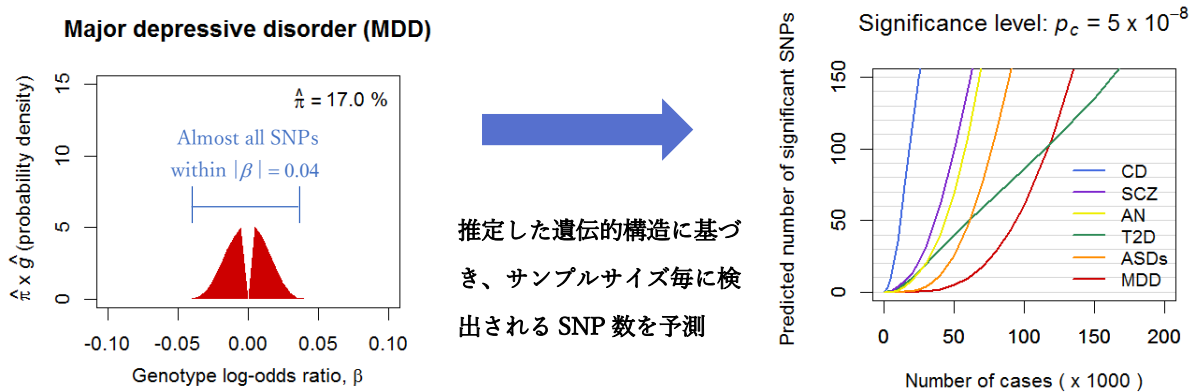


大うつ病 GWAS が成功するために必要なサンプルサイズ

大うつ病性障害 (major depressive disorder; MDD) は、患者や家族の生活の質を著しく低下させ、その生涯有病率は約 15%にもする、社会経済的損失が大きい疾患として知られています。一方で、MDD は中程度の遺伝性 (遺伝率 = 31–42%) が確認され遺伝学的な発症機序の理解が期待できる複雑疾患でもあります。ゲノムワイド関連解析 (genome-wide association study; GWAS) は、ゲノム全体の一塩基多型 (single nucleotide polymorphism; SNP) と形質との関連を網羅的に調べる方法であり、これまで多くの複雑疾患の感受性変異を同定してきました。MDD においても複数の GWAS 研究がなされてきましたが、約 1 万例にも及ぶ症例を用いても、ひとつの疾患関連 SNP も検出できず、GWAS 研究の最も難しい複雑疾患であることが分かってきました。

最近、我々は複雑疾患の GWAS 要約データを用いて、その背後にある多遺伝因子構造 (疾患関連 SNP の割合やそれらの効果サイズの分布) を柔軟に推定する「セミパラメトリック階層混合モデル」 (Semi-parametric hierarchical mixture model; SP-HMM) を開発しました [1]。推定した多遺伝因子構造の下では、任意のサンプルサイズに対して統計的有意水準 (例えばゲノムワイド有意水準 $< 5 \times 10^{-8}$) を満たす疾患関連 SNP の検出数の推定が可能となります。そこで、本研究では、SP-HMM を用いて MDD の多遺伝因子構造の特徴を明らかにしつつ、今後サンプルサイズを増やしていったときに疾患関連 SNP の検出に成功しうるのかを調べました。

その結果、MDD は、他の疾患と比べ、疾患関連 SNP の割合 (π) が小さくそれらのリスクアリのオッズ比は非常に小さい (ほとんどのリスクアリのオッズ比はわずか ≤ 1.04 に収まる) ことが明らかになりました。そして、疾患関連 SNP の検出効率は、際立って小さいことが分かりました。それでも、約 10 万ずつの症例と対照を用いれば (これは将来的には到達可能なサンプル数と考えられます)、約 50 個もの関連 SNP が検出できるという推定を得ることができました。本解析は、MDD 以外の複雑疾患の GWAS についても、疾患関連 SNP 特定に必要なサンプルサイズを予測するために有用な方法であると考えられます。



CD:クローン病; SCZ:統合失調症; AN:神経性無食欲症; T2D: II 型糖尿病; ASDs:自閉症スペクトラム障害

1. Nishino J, Kochi Y, Shigemizu D, Kato M, Ikari K, Ochi H, Noma H, Matsui K, Morizono T, Boroevich K.A, Tsunoda T, Matsui S. (2016). Empirical Bayes estimation of semi-parametric hierarchical mixture models for unbiased characterization of polygenic disease architectures. bioRxiv doi.org/10.1101/080945