Genetic bases of schizophrenia phenotypes: abridged secondary publication

KCY Wong, ST Rao, S Zhi, PBM Leung, BKW Lee, EFC Cheung, RCK Chan, KKY Ho, KSY Hung, TCK Hui, PC Sham, SSY Lui, HC So *

KEY MESSAGE

Understanding the genetic architecture of schizophrenia phenotypes, shared pathophysiology, and genetic and clinical risk factors associated with disease severity may inform patient classification and guide risk stratification for treatment prioritisation.

Hong Kong Med J 2026;32(Suppl 1):S22-3
HMRF project number: 07180376

[1] KCY Wong, [2] ST Rao, [2] S Zhi, [3] PBM Leung, [3] BKW Lee, [4] EFC Cheung, [3] RCK Chan, [5] KKY Ho, [6] KSY Hung, [3] TCK Hui, [3] PC Sham, [3] SSY Lui, [1] HC So

PC Sham, SSY Lui, and HC So contributed equally

[1] *School of Biomedical Sciences, The Chinese University of Hong Kong, Hong Kong SAR, China*
[2] *School of Medical Technology and Engineering, Fujian Medical University, Fuzhou, China*
[3] *Department of Psychiatry, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China*
[4] *Queen Elizabeth Hospital , Hong Kong SAR, China*
[5] *Castle Peak Hospital, Hong Kong SAR, China*
[6] *Hong Kong Hospital Authority, Hong Kong SAR, China*

* Principal applicant and corresponding author: hcso@cuhk.edu.hk

Introduction

Schizophrenia (SCZ) is highly heritable and characterised by wide variability in clinical presentation and prognosis.[1] Genetic factors contribute substantially to this heterogeneity.[2] Understanding the genetic basis is essential for improving diagnosis and treatment. Despite advances in genome-wide association studies (GWAS), the complex aetiology of SCZ remains unclear.[3-5] Most GWAS have focused on the diagnosis itself, without detailed investigation of symptoms, course, and outcomes. Large-scale GWAS involving diverse clinical phenotypes and prognosis in Chinese populations are scarce, limiting identification of specific risk alleles and genes. Furthermore, the extent to which different clinical features and outcomes in SCZ share genetic underpinnings with the core disorder and with other psychiatric conditions remains largely unexplored. This study aimed to investigate the genetic basis of SCZ phenotypes in the Chinese population.

Methods

We analysed a variety of phenotypes, including age at onset, history of self-harm and aggression, and number of psychiatric hospitalisations since onset, using GWAS (with meta-analyses across Chinese and European cohorts), transcriptome-wide association studies, polygenic risk score analysis, gene-set and pathway enrichment analyses, machine learning modelling with Shapley value analysis, and gene-based analytical tools such as MAGMA and MetaXcan.

Results

GWAS identified genome-wide significant single nucleotide polymorphisms (SNPs) associated with age at onset and the number of hospitalisations (both $P<5\times10^{-8}$). A meta-analysis combining Chinese and European cohorts re-identified the same significant SNPs associated with age at onset.

Polygenic risk score analyses revealed significant genetic associations between SCZ phenotypes and other psychiatric disorders. Gene-based analyses and transcriptome-wide association studies implicated specific genes associated with the immune system and brain tissues. Our machine learning model achieved a moderate level of discriminative performance, comparable to that in a previous study,[6] with several predictors identified by Shapley value analysis.

Discussion

The identification of significant SNPs, genes, and biological pathways highlights the importance of specific molecular mechanisms and brain regions in SCZ pathogenesis. These insights may have important implications for risk stratification, such as the early identification of high-risk individuals and guidance for differential diagnosis in ambiguous cases. By elucidating the genetic overlap between SCZ phenotypes and other psychiatric disorders or traits, the present study provides a deeper understanding of the genetic architecture of SCZ and its shared pathophysiology with related phenotypes.

We developed a machine learning model with moderate accuracy to assess the severity of

SCZ phenotypes. This model provides a proof of concept for how combined genetic and clinical risk factors can improve risk stratification and disease subtyping. Such model is particularly relevant for patients with poorer prognosis, who may benefit from prioritised interventions, such as more intensive psychosocial and rehabilitative treatments. Moreover, the analytical frameworks developed to integrate genetic and clinical data in SCZ are broadly applicable to other disorders.

The present study has several limitations. First, the modest sample size may limit statistical power and weaken the reliability of GWAS and polygenic risk score analysis results. The prioritised genetic variants or genes should therefore be considered tentative and require further confirmation in future studies. Second, the use of mismatched ethnic groups in some polygenic risk score estimations may reduce statistical power. Finally, the machine learning model may be biased towards the characteristics of this specific dataset; its generalisability warrants further study. Additional validation in larger and more ethnically diverse samples is required before clinical translation.

Candidate genes identified for clinical phenotypes may serve as useful targets for follow-up functional studies and investigation in other psychiatric disorders. Further research is warranted to identify causal risk factors for SCZ outcomes using Mendelian randomisation, with genetic variants as instruments. Findings on shared genetic bases between SCZ phenotypes and other psychiatric disorders or traits highlight the potential of applying genetics and polygenic risk scores in differential diagnosis. Accurate diagnosis is critical in psychiatry, as treatments and prognosis can differ substantially.

Conclusion

Our study deepens the understanding of the genetic architecture of SCZ phenotypes, shared pathophysiology, and the genetic and clinical risk factors associated with disease severity. Our findings may inform patient classification and guide risk stratification for treatment prioritisation. However, rigorous validation in larger, ethnically diverse cohorts is needed.

Funding

This study was supported by the Health and Medical Research Fund, Health Bureau, Hong Kong SAR Government (#07180376). The full report is available from the Health and Medical Research Fund website (https://rfs1.healthbureau.gov.hk).

Disclosure

The results of this research have been previously published in:

1. Rao ST, Wong KCY, Zhi S, et al. Unraveling genetic variants underlying schizophrenia phenotypes: an original GWAS in Hong Kong Chinese with cross-ethnic meta-analysis and predictive modeling [preprint]. medRxiv 2025:2025.12.24.25342953.
2. Tubbs JD, Leung PBM, Zhong Y, et al. Pathway-specific polygenic scores improve cross-ancestry prediction of psychosis and clinical outcomes [preprint]. medRxiv 2023:2023.09.01.23294957.
3. Zhan N, Sham PC, So HC, Lui SSY. The genetic basis of onset age in schizophrenia: evidence and models. Front Genet 2023;14:1163361.

Acknowledgements

We thank the EASY team staff at Castle Peak Hospital, and Dr Karen KY Ho, Dr Karen SY Hung, and Mr Tomy CK Hui from Department of Psychiatry, The University of Hong Kong for their valuable contributions.

References

1. Global Burden of Disease Study 2013 Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2015;386:743-800.
2. Fanous AH, Kendler KS. Genetic heterogeneity, modifier genes, and quantitative phenotypes in psychiatric illness: searching for a framework. Mol Psychiatry 2005;10:6-13.
3. Bani-Fatemi A, Graff A, Zai C, Strauss J, De Luca V. GWAS analysis of suicide attempt in schizophrenia: main genetic effect and interaction with early life trauma. Neurosci Lett 2016;622:102-6.
4. Edwards AC, Bigdeli TB, Docherty AR, et al. Meta-analysis of positive and negative symptoms reveals schizophrenia modifier genes. Schizophr Bull 2016;42:279-87.
5. Fanous AH, Zhou B, Aggen SH, et al. Genome-wide association study of clinical dimensions of schizophrenia: polygenic effect on disorganized symptoms. Am J Psychiatry 2012;169:1309-17.
6. Bracher-Smith M, Rees E, Menzies G, et al. Machine learning for prediction of schizophrenia using genetic and demographic factors in the UK biobank. Schizophr Res 2022;246:156-64.