

Research Article

Title

Genetic overlap between epilepsy and schizophrenia: evidence from cross phenotype analysis in Hong Kong Chinese population¹

Hongsheng Gui 0000-0002-0460-7446 0000-0002-0460-7446^{1,2*}, Miaoxin Li³, Pak C. Sham^{1,3,4}, Larry Baum³, International League Against Epilepsy Consortium on Complex Epilepsies[&], Patrick Kwan^{5,6}, Stacey S. Cherny^{1,3,4,*}

¹ Center for Genomic Sciences, The University of Hong Kong, Hong Kong, SAR, China; ² Center for Health Policy and Health Research Service, Henry Ford Health System, Detroit, Michigan, United States of America; ³ Department of Psychiatry, The University of Hong Kong, Hong Kong, SAR, China; ⁴ The State Key Laboratory of Brain and Cognitive Sciences, The University of Hong Kong, Hong Kong, SAR, China; ⁵ Department of Medicine & Therapeutics, the Chinese University of Hong Kong, Hong Kong, SAR, China; ⁶ Department of Neurology, Royal Melbourne Hospital, Melbourne, Australia

* Corresponding author contact information: Hongsheng Gui (allenkuei@gmail.com), and Stacey Cherny (cherny@hku.hk) at Room 1-05G, 1/F, The Hong Kong Jockey Club Building for Interdisciplinary Research, University of Hong Kong, 5 Sassoon Road, Pokfulam, Hong Kong SAR, China; phone number: +85228315079; fax number: +85228185653. & Members listed at end of paper.

¹ This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi:[10.1002/ajmgb.32607](https://doi.org/10.1002/ajmgb.32607)

Abstracts

Epilepsy and schizophrenia are common and typical neurological or mental illness respectively, and sometimes they comorbid in the same patients, however the underlying genetic relationship between the two brain diseases is still not fully understood. To investigate the possible genetic contribution to their comorbidity, we performed polygenic risk score (PRS) analyses and genetic correlation estimation so as to identify the overall genetic overlap between the two diseases. The global schizophrenia PRS is strongly associated with schizophrenia phenotype in Hong Kong population (odds ratio=1.7, $p=2.26E-16$), and focal epilepsy PRS is moderately associated with epilepsy phenotype in Hong Kong population (odds ratio=1.14, $p=0.013$). However the disease-specific PRS can only predict its own well-matched phenotype but not the other ones ($p > 0.05$). This pattern is further supported by non-significant pairwise genetic correlation and insufficient statistical power for PRS association from the cross-phenotype analyses. Our study reveals there's limited shared genetic aetiology between schizophrenia and epilepsy, and thus supports a model of shared environmental factors to explain the comorbidity between the two phenotypes.

Keywords: genetic sharing; comorbidity; GWAS; polygenic risk score

Introduction

Epilepsy and schizophrenia are two common brain-involved disorders. Patients with epilepsy are at increased risk of developing psychosis compared with the general population [Chen et al., 2016; Tellez-Zenteno et al., 2007; Torta and Keller 1999; Vuilleumier and Jallon 1998] and some forms of seizures may mimic the symptoms of psychosis [Sethi and Emery 2013; Tucker 1998]; therefore better understanding of epilepsy might also provide insights on the understanding of mental illness like schizophrenia, and vice versa. One example is from brain image studies, which revealed hippocampus volume is probably related to both epilepsy and schizophrenia [Jack 1994; Nelson et al., 1997]. Another example comes from a family-based study that gave evidence of 2-fold increase in the risk of developing psychosis for individuals with a parental history of epilepsy [Clarke et al., 2012]. Moreover, comorbidity analysis from previous epidemiology study in Chinese [Chang et al., 2011] found higher incidence of epilepsy in schizophrenia cohort than non-schizophrenia cohort (6.99 vs. 1.19 per 1000 person-years) and also higher incidence of schizophrenia in epilepsy cohort than

non-epilepsy cohort (3.53 vs. 0.46 per 1000 person-years), hence showed they may relate to each other at disease susceptibility.

However, from the perspective of genetics, few susceptibility loci have been found to predispose to both two diseases [Welter et al., 2014]. Moreover, genetic overlaps between them are not clear due to different hypothesis in previous candidate-gene based research. Findings from previous genome-wide association studies (GWAS) for epilepsy and schizophrenia that are stored in GWAS Catalog (<https://www.ebi.ac.uk/gwas/home>) share very few hits for both. This can be ascribed to limited power of univariate SNP analysis, given GWAS variants are mostly with small effect sizes on the phenotypes [Park et al., 2010]. Nevertheless, GWAS provide good resources for secondary analyses that aim to leverage the power of multiple phenotypes, or establish the genetic relationship between them [Bulik-Sullivan et al., 2015a; Solovieff et al., 2013]. Among different data mining approaches, PRS analysis is widely used to search for shared genetic aetiology between two datasets on the same phenotype or different phenotypes [Chatterjee et al., 2016; Purcell et al., 2009]. Regarding to epilepsy and schizophrenia, large-scale GWAS meta-analyses have already been reported for them respectively [International League Against Epilepsy Consortium on Complex Epilepsies. Electronic address 2014; Ripke et al., 2014] and can provide necessary training dataset for PRS construction. In addition, our previous GWAS reports on both schizophrenia and epilepsy can serve as independent target dataset for testing PRS association [Guo et al., 2012; Wong et al., 2014].

In this study, we systematically estimated the overall genetic sharing by using publicly available summary statistics and in-house genome-wide raw genotypes for both epilepsy and schizophrenia. We focused mainly on PRS analysis in our Chinese population, together with genetic correlation analysis between different datasets, so as to establish the genotype-phenotype relationship across these two diseases.

Methods

Data collection and preprocessing

For training datasets, we collected summary statistics (SNP ID, alleles, p -values, odds ratio (OR), and standard errors) from GWAS meta-analyses previously

reported by International League Against Epilepsy (ILAE) consortium on complex epilepsies [International League Against Epilepsy Consortium on Complex Epilepsies. Electronic address 2014] and Psychiatric genomics consortium (PGC) [Ripke et al., 2014] respectively. The latest meta-analysis result on schizophrenia for 49 Caucasian population cohorts (in total 33,640 cases and 43,456 controls) was downloaded from PGC public repository (<http://www.med.unc.edu/pgc/>). Regarding to epilepsy, we re-performed fixed-effect meta-analyses for all epilepsies (including both genetic generalized epilepsy (GGE) and focal epilepsy) using Caucasian samples (7019 cases and 19935 controls) and for focal epilepsy separately (4050 cases and 19935 controls), since Hong Kong cohort was comprised of focal epilepsy patients only.

For target datasets, we retrieved raw genotype data from two previous GWAS done in our team, for epilepsy or schizophrenia in Hong Kong (HK) Chinese population respectively [Guo et al., 2012; Wong et al., 2014]. Samples from both GWAS were genotyped on the Human610-Quad BeadChip (Illumina, Inc., San Diego, CA). Both datasets have been imputed to 1000 Genome reference panel using the same protocol. Moreover, to eliminate the influence of different controls on polygenic score estimation for two target datasets, we kept only 1661 shared controls between them. For HK GWAS datasets after standard quality control, we estimated the population structure by principal component (PC) analysis using linkage disequilibrium (LD) pruned SNPs [Patterson et al., 2006]; then we generated univariate SNP-level summary statistics by running PLINK logistic regression between binary phenotype and additive minor allele copies (0, 1 or 2) [Purcell et al., 2007], while adjusting for significant PCs determined by R package 'EigenCorr' [Lee et al., 2011]. To reduce mismatched markers between datasets from different resources, we restricted genetic markers to HapMap3 variants only. A python script in LD Score Regression (LDSC) package (<https://github.com/bulik/ldsc>) was used to filter out unqualified variants from either training or target datasets. **Table-1** lists the final data we used in this study.

PRS association analysis

For each of two target datasets from Hong Kong epilepsy GWAS (HKEPI) and Hong Kong schizophrenia GWAS (HKSCZ), we calculated its polygenic risk scores using all epilepsies meta-GWAS (EPIall), focal epilepsy meta-GWAS (EPIfocal) or schizophrenia

meta-GWAS (SCZ) as training dataset (Table 1). In brief, independent SNPs were generated by LD clumping ($r^2 < 0.1$) in target dataset and then summed up after weighting each by its effect size in training dataset. A logistic regression was then used to test the association between each PRS and the corresponding phenotype (binary status of disease), with adjustment of the same principal components from univariate SNP analysis. To select different groups of SNPs in PRS, we switched inclusion p -value threshold from 0 to 0.5 by 0.005 increment and chose the best-fit model according to its variance explained (Nagelkerke's pseudo R^2). PRSice was used to perform the calculation, model fitting and visualization [Euesden et al., 2015]. In order to present a more meaningful effect size, the raw PRSs were standardized to have distribution with mean 0 and variance 1. ORs relative to one standard deviation increase were computed using beta estimates from logistic regression model newly fitted for standardized PRS. This transformation and modeling was done manually in R program (<https://www.r-project.org/>).

Heritability and genetic correlation

LDSC approach [Bulik-Sullivan et al., 2015a] was used to calculate heritability at liability scale for each dataset (SCZ, EPIall, EPIfocal, HKSCZ and HKEPI), also genetic co-variance and correlation between each pair of the five datasets. Latest python script `ldsc.py` (<https://github.com/bulik/ldsc>, May 2017) was applied for all calculations, with HapMap Caucasian or East Asian reference panel to adjust for LD pattern between genetic markers.

Power estimation

We used 'polygenescore' function to calculate the power of PRS association analyses conducted in this study [Dudbridge 2013], with R code downloaded from <https://sites.google.com/site/fdudbridge/software/>. For each pair involving one training dataset and one target dataset, disease prevalences (K_1 , K_2), sample size (n_1 , n_2) and sampling ratios (P_1 , P_2) were from Table 1, while heritability in training dataset (σ_1^2) and genetic co-variance between two datasets (σ_{12}) were acquired by calculation in above section. Numbers of independent markers (m) were estimated by LD clumping at $r^2 < 0.1$ for GWAS genotypes from target dataset, and proportion of markers with no effect (π_0) was defined as percentage of independent SNPs with p -value > 0.05 in summary statistics from training dataset. Type I error rate (α)

was set at 0.001.

Results

Genomic inflation factors (λ) in training datasets are 1.61 for SCZ, 0.99 for EPIall, and 0.92 for EPIfocal; in target dataset, λ is 1.02 and 1.04 for HKSCZ and HKEPI respectively after adjusting for two significant PCs. The bar charts (Figure 1) show the model fitting and associations of PRS from various p -value thresholds on the different datasets. Firstly, schizophrenia PRSs are strongly associated with the status of schizophrenia in HK population (Figure 1A), with variance explained (pseudo R^2) at 5.1%, and p -value at $2.26E^{-16}$ for the best-fit model; however, SCZ PRS does not predict HK epilepsy status at all (Figure 1B) since all p -values are above 0.05. Secondly, all epilepsies PRSs are not associated with either HKSCZ (Figure 1C) or HKEPI (Figure 1D). Thirdly, though focal epilepsy PRSs don't predict the status of HKSCZ either (Figure 1E), the best-fit PRS is moderately associated with HKEPI (variance explained=0.4%, $p=0.013$) (Figure 1F). Furthermore, stratified analysis on HKEPI reveals EPIfocal PRS associates mostly with HK focal epilepsy patients documented hippocampal sclerosis lesion (variance explained=1.9%, $p=0.001$), but not with other sub-phenotypes (Figure S1).

With selection of three fixed thresholds (0.005, 0.05 and 0.5) and the one from best-fit model, odds ratios computed using standardized PRS are shown in Table 2. SCZ PRS consistently increases the risk of schizophrenia in Hong Kong population at all different thresholds (OR >1.6, $p<1.0E-13$), and focal epilepsy PRS only slightly increases the risk of epilepsy in Hong Kong (OR=1.14, $p=0.013$). No pairs crossing two different diseases are found with statistical evidence of PRS association, indicating there is limited genetic overlapping between epilepsy and schizophrenia.

Table 3 gives the estimated heritability for each dataset and genetic correlation for each pair of datasets, using summary statistics from common SNPs. Only two pairs are found with significant genetic correlations (SCZ and HKSCZ, $R^2=0.39$, $p=4.9E-10$; EPIall and EPIfocal, $R^2=0.91$, $p=3.7E-275$). Similar to PRS association, we don't find any evidence of shared genetic aetiology between schizophrenia and epilepsy, even with up-to-date largest meta-analysis datasets (SCZ and EPIall, $R^2=0.08$, $p=0.164$).

Finally, we plugged epidemiological parameters in Table 1 and LDSC calculated genetic parameters in Table 3 into the function of power calculation for PRS

association in Table 2. As expected, current training dataset from schizophrenia is well powered for predicting HK SCZ at all different thresholds (Table 4). The comparatively lower power (<0.5) of predicting HK EPI by SCZ PRS reflects limited genetic correlation between epilepsy and schizophrenia. All other power estimates are under 0.1 (Table 4). This may be due to much smaller sample size and SNP-based heritability from training dataset (EPI_{all} or EPI_{focal}), or inaccurate genetic covariance (with wide standard error) between two datasets with relatively low sample size in either one or both.

Discussion

Epilepsy and schizophrenia are two common diseases belonging to the same broad category of brain disorders, which cause huge economic burden for the patient, his/her families and the society. Our study reveals the potential of using PRS to predict the same phenotype (especially on schizophrenia); in addition it suggests no significant amount of genetic overlap may exist between epilepsy and schizophrenia.

Secondary analyses on GWAS dataset are still providing new insights for complex phenotypes [Barabasi et al., 2011; Bulik-Sullivan et al., 2015a; Chatterjee et al., 2016]. The more advanced data mining approaches can help to reveal hidden genetic signals for one phenotype or multiple phenotypes together [Cantor et al., 2010; Yang et al., 2012]. We previously performed multilevel GWAS analysis to search for overlapping functional units (genic SNPs, genes or gene-sets) across different complex diseases [Gui et al., 2017], but found no significant sharing between epilepsy and schizophrenia (data not shown). To further examine the genetic relationship between schizophrenia and epilepsy, we resorted to PRS analysis and genetic correlation analysis in this study. These two approaches have been independently established and applied in a few complex diseases or traits before [Bulik-Sullivan et al., 2015a; Chatterjee et al., 2016; Purcell et al., 2009]. In our study the results from them agree mostly with each other, as both support the close relationship between global schizophrenia and Hong Kong schizophrenia, but not across epilepsy and schizophrenia. Given the nominal effect of SCZ PRS on our epilepsy patients ($OR > 1$, $P < 0.1$) and $<80\%$ statistical power across the two phenotypes, we cannot completely rule out the possibility of elevating risk of epilepsy through schizophrenia. Nevertheless, the lack of genetic overlap between these two disorders is further supported by a recent finding that reveals substantial genetic

difference between neurological and psychiatric disorders [Anttila et al., 2016]. The genetic correlation coefficients between different kinds of epilepsy and schizophrenia are all reported at near-zero level. Though an earlier study using data from PGC1 and ILAE meta-analyses identified significant correlation between all epilepsies and schizophrenia [Vonberg and Bigdeli 2016], we did not verify it when excluding Hong Kong cohort (schizophrenia patients served as controls to epilepsy patients in GWAS) from ILAE meta-analysis.

In comparison to clear relationship within schizophrenia datasets or across schizophrenia and epilepsy datasets, the pattern within epilepsy datasets is more vague and arbitrary. From PRS analysis, we identified moderate association of EPIfocal risk (instead of EPIall risk) on Hong Kong focal epilepsy patients (especially for the sub-phenotype characterized by hippocampal sclerosis). However, there is no significant genetic correlation found between epilepsy training and target datasets. A few factors may explain the inefficiency or discrepancy observed for epilepsy. First of all, epilepsy training dataset has much lower sample size than schizophrenia, hence can identify limited polygenicity among common SNPs [Bulik-Sullivan et al., 2015b]; accordingly, the explained heritability by the same SNP array is much lower than that for schizophrenia. Secondly, both epilepsy training dataset and target dataset have more heterogeneous phenotype than that for schizophrenia. Thirdly, the LD pattern in training dataset (Caucasian population) is different from targeting dataset (Asian population), hence may cause different extent of power reduction or bias estimate from either PRS prediction [Martin et al., 2017] or genetic correlation analysis [Bulik-Sullivan et al., 2015a]. All these factors contribute to the small power estimates (Table 4) for phenotype pairs involving epilepsy dataset and highlight the limitation of current study. Specifically, we have observed better model fitting (variance explained from 2.8% to 5.1%) for SCZ PRS on HKSCZ when switching from PGC1 to PGC2. Therefore improvement from these factors, either a magnitude increase of sample size, or homogenous phenotype selection, or well-matched population between training and target dataset, can be helpful to improve PRS association and risk prediction for epilepsy in later studies.

The lack of genetic sharing at genome-wide scale should redirect the effort to explain the comorbidity of epilepsy and schizophrenia to other potential mechanisms including environmental factors [Rudzinski and Meador 2013]. Neurological deficits at birth (including mental retardation and cerebral palsy) can

increase the susceptibility to both diseases [Ottman et al., 1996]. Antecedent central nervous system injury (including stroke, trauma and brain infection) may lead to hippocampal volume deficient and then cause the onset of epilepsy and schizophrenia [Schmitt et al., 2014]. Other factors, like stress during development, depression and anti-epileptic drug effects on mental state, are all important contributor to the comorbidity [Agrawal and Govender 2011; Barnes and Paolicchi 2008]. Deeper understanding on the direct pathway from these environmental risk factors to two phenotypes has to await future studies.

International League Against Epilepsy Consortium on Complex Epilepsies

Richard J L Anney, Andreja Avbersek, David Balding, Larry Baum, Felicitas Becker, Samuel F Berkovic, Jonathan P Bradfield, Lawrence C Brody, Russell J Buono, Claudia B Catarino, Gianpiero L Cavalleri, Stacey S Cherny, Krishna Chinthapalli, Alison J Coffey, Alastair Compston, Patrick Cossette, Gerrit-Jan de Haan, Peter De Jonghe, Carolien G F de Kovel, Norman Delanty, Chantal Depondt, Dennis J Dlugos, Colin P Doherty, Christian E Elger, Thomas N Ferraro, Martha Feucht, Andre Franke, Jacqueline French, Verena Gaus, David B Goldstein, Hongsheng Gui, Youling Guo, Hakon Hakonarson, Kerstin Hallmann, Erin L Heinzen, Ingo Helbig, Helle Hjalgrim, Margaret Jackson, Jennifer Jamnadas-Khoda, Dieter Janz, Michael R Johnson, Reetta Kälviäinen, Anne-Mari Kantanen, Dalia Kasperavičiūtė, Dorothee Kasteleijn-Nolst Trenite, Bobby P C Koeleman, Wolfram S Kunz, Patrick Kwan, Yu Lung Lau, Anna-Elina Lehesjoki, Holger Lerche, Costin Leu, Wolfgang Lieb, Dick Lindhout, Warren Lo, Daniel H Lowenstein, Alberto Malovini, Anthony G Marson, Mark McCormack, James L Mills, Martina Moerzinger, Rikke S Møller, Anne M Molloy, Hiltrud Muhle, Mark Newton, Ping-Wing Ng, Markus M Nöthen, Peter Nürnberg, Terence J O'Brien, Karen L Oliver, Aarno Palotie, Faith Pangilinan, Katharina Pernhorst, Slave Petrovski, Michael Privitera, Rodney Radtke, Philipp S Reif, Felix Rosenow, Ann-Kathrin Ruppert, Thomas Sander, Theresa Scattergood, Steven Schachter, Christoph Schankin, Ingrid E Scheffer, Bettina Schmitz, Susanne Schoch, Pak C Sham, Sanjay Sisodiya, David F Smith, Philip E Smith, Doug Speed, Michael R Sperling, Michael Steffens, Ulrich Stephani, Pasquale Striano, Hans Stroink, Rainer Surges, K Meng Tan, The KORA study group, G Neil Thomas, Marian Todaro, Anna Tostevin, Rossana Tozzi, Holger Trucks, Frank Visscher, Sarah von Spiczak, Nicole M Walley, Yvonne G Weber, Zhi Wei, Christopher Whelan, Wanling Yang, Federico Zara, and Fritz Zimprich.

Acknowledgements

This study is supported by Small Project Funding from the University of Hong Kong (SPF 201409176131 to HS Gui). We would like to thank Dr. Timothy Mak for technical advice on running PRISice, and Psychiatric Genomic Consortium for releasing their schizophrenia GWAS meta-analysis result available online.

Competing interests

None declared.

References

Agrawal N, Govender S. 2011. Epilepsy and neuropsychiatric comorbidities. *Advances in*

- Psychiatric Treatment 17:44-53.
- Anttila V, Bulik-Sullivan B, Finucane HK, Bras J, Duncan L, Escott-Price V, Falcone G, Gormley P, Malik R, Patsopoulos N, Ripke S, Walters R, Wei Z, Yu D, Lee P, Breen G, Bulik C, Daly M, Dichgans M, Faraone S, Guerreiro R, Holmans P, Kendler K, Koeleman B, Mathews C, Scharf J, Sklar P, Williams J, Wood N, Cotsapas C, Palotie A, Smoller 0000-0002-0381-6334 J, Sullivan P, Rosand J, Corvin A, Neale B. 2016. Analysis of shared heritability in common disorders of the brain. *bioRxiv*.
- Barabasi AL, Gulbahce N, Loscalzo J. 2011. Network medicine: a network-based approach to human disease. *Nat Rev Genet* 12(1):56-68.
- Barnes GN, Paolicchi JM. 2008. Neuropsychiatric comorbidities in childhood absence epilepsy. *Nature clinical practice Neurology* 4(12):650-651.
- Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR, ReproGen C, Psychiatric Genomics C, Genetic Consortium for Anorexia Nervosa of the Wellcome Trust Case Control C, Duncan L, Perry JR, Patterson N, Robinson EB, Daly MJ, Price AL, Neale BM. 2015a. An atlas of genetic correlations across human diseases and traits. *Nat Genet* 47(11):1236-1241.
- Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics C, Patterson N, Daly MJ, Price AL, Neale BM. 2015b. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* 47(3):291-295.
- Cantor RM, Lange K, Sinsheimer JS. 2010. Prioritizing GWAS results: A review of statistical methods and recommendations for their application. *Am J Hum Genet* 86(1):6-22.
- Chang YT, Chen PC, Tsai IJ, Sung FC, Chin ZN, Kuo HT, Tsai CH, Chou IC. 2011. Bidirectional relation between schizophrenia and epilepsy: A population-based retrospective cohort study. *Epilepsia* 52(11):2036-2042.
- Chatterjee N, Shi JX, Garcia-Closas M. 2016. Developing and evaluating polygenic risk prediction models for stratified disease prevention. *Nat Rev Genet* 17(7):392-406.
- Chen Z, Lusicic A, O'Brien TJ, Velakoulis D, Adams SJ, Kwan P. 2016. Psychotic disorders induced by antiepileptic drugs in people with epilepsy. *Brain : a journal of neurology* 139(Pt 10):2668-2678.
- Clarke MC, Tanskanen A, Huttunen MO, Clancy M, Cotter DR, Cannon M. 2012. Evidence for shared susceptibility to epilepsy and psychosis: a population-based family study. *Biological psychiatry* 71(9):836-839.
- Dudbridge F. 2013. Power and Predictive Accuracy of Polygenic Risk Scores. *Plos Genet* 9(3).
- Euesden J, Lewis CM, O'Reilly PF. 2015. PRSice: Polygenic Risk Score software. *Bioinformatics* 31(9):1466-1468.
- Gui H, Kwan JS, Sham PC, Cherny SS, Li M. 2017. Sharing of Genes and Pathways Across Complex Phenotypes: A Multilevel Genome-Wide Analysis. *Genetics*.
- Guo Y, Baum LW, Sham PC, Wong V, Ng PW, Lui CH, Sin NC, Tsoi TH, Tang CS, Kwan JS, Yip BH, Xiao SM, Thomas GN, Lau YL, Yang W, Cherny SS, Kwan P. 2012. Two-stage genome-wide association study identifies variants in CAMSAP1L1 as susceptibility loci for epilepsy in Chinese. *Hum Mol Genet* 21(5):1184-1189.
- International League Against Epilepsy Consortium on Complex Epilepsies. Electronic address e-aeua. 2014. Genetic determinants of common epilepsies: a meta-analysis of genome-wide association studies. *Lancet Neurol* 13(9):893-903.
- Jack CR. 1994. MRI-Based Hippocampal Volume Measurements in Epilepsy. *Epilepsia* 35:S21-S29.
- Lee S, Wright FA, Zou F. 2011. Control of population stratification by correlation-selected principal components. *Biometrics* 67(3):967-974.
- Martin AR, Gignoux CR, Walters RK, Wojcik GL, Neale BM, Gravel S, Daly MJ, Bustamante CD, Kenny EE. 2017. Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. *Am J Hum Genet* 100(4):635-649.
- Nelson MD, Saykin AJ, Flashman LA, Riordan HJ. 1997. Hippocampal volume reduction in schizophrenia as assessed by magnetic resonance imaging: A meta-analytic study.

Schizophr Res 24(1-2):153-153.

- Ottman R, Annegers JF, Risch N, Hauser WA, Susser M. 1996. Relations of genetic and environmental factors in the etiology of epilepsy. *Annals of neurology* 39(4):442-449.
- Park JH, Wacholder S, Gail MH, Peters U, Jacobs KB, Chanock SJ, Chatterjee N. 2010. Estimation of effect size distribution from genome-wide association studies and implications for future discoveries. *Nat Genet* 42(7):570-575.
- Patterson N, Price AL, Reich D. 2006. Population structure and eigenanalysis. *Plos Genet* 2(12):e190.
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81(3):559-575.
- Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF, Sklar P, Ruderfer DM, McQuillin A, Morris DW, O'Dushlaine CT, Corvin A, Holmans PA, Macgregor S, Gurling H, Blackwood DHR, Corvin A, Craddock NJ, Gill M, Hultman CM, Kirov GK, Lichtenstein P, Muir WJ, Owen MJ, Pato CN, Scolnick EM, St Clair D, Craddock NJ, Holmans PA, Williams NM, Georgieva L, Nikolov I, Norton N, Williams H, Toncheva D, Milanova V, Hultman CM, Lichtenstein P, Thelander EF, Sullivan P, Kenny E, Quinn EM, Gill M, Corvin A, Choudhury K, Datta S, Pimm J, Thirumalai S, Puri V, Krasucki R, Lawrence J, Quedsted D, Bass N, Crombie C, Fraser G, Kuan SL, Walker N, Blackwood DHR, Muir WJ, McGhee KA, Pickard B, Malloy P, Maclean AW, Van Beck M, Wray NR, Macgregor S, Visscher PM, Pato MT, Medeiros H, Middleton F, Carvalho C, Morley C, Fanous A, Conti D, Knowles JA, Ferreira CP, Macedo A, Azevedo MH, Kirby AN, Ferreira MAR, Daly MJ, Chambert K, Kuruvilla F, Gabriel SB, Ardlie K, Moran JL, Daly MJ, Scolnick EM. 2009. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 460(7256):748-752.
- Ripke S, Neale BM, Corvin A, Walters JTR, Farh KH, Holmans PA, Lee P, Bulik-Sullivan B, Collier DA, Huang HL, Pers TH, Agartz I, Agerbo E, Albus M, Alexander M, Amin F, Bacanu SA, Begemann M, Belliveau RA, Bene J, Bergen SE, Bevilacqua E, Bigdeli TB, Black DW, Bruggeman R, Buccola NG, Buckner RL, Byerley W, Cahn W, Cai GQ, Champion D, Cantor RM, Carr VJ, Carrera N, Catts SV, Chambert KD, Chan RCK, Chen RYL, Chen EYH, Cheng W, Cheung EFC, Chong SA, Cloninger CR, Cohen D, Cohen N, Cormican P, Craddock N, Crowley JJ, Curtis D, Davidson M, Davis KL, Degenhardt F, Del Favero J, Demontis D, Dikeos D, Dinan T, Djurovic S, Donohoe G, Drapeau E, Duan J, Dudbridge F, Durmishi N, Eichhammer P, Eriksson J, Escott-Price V, Essioux L, Fanous AH, Farrell MS, Frank J, Franke L, Freedman R, Freimer NB, Friedl M, Friedman JI, Fromer M, Genovese G, Georgieva L, Giegling I, Giusti-Rodriguez P, Godard S, Goldstein JI, Golimbet V, Gopal S, Gratten J, de Haan L, Hammer C, Hamshere ML, Hansen M, Hansen T, Haroutunian V, Hartmann AM, Henskens FA, Herms S, Hirschhorn JN, Hoffmann P, Hofman A, Hollegaard MV, Hougaard DM, Ikeda M, Joa I, Julia A, Kahn RS, Kalaydjieva L, Karachanak-Yankova S, Karjalainen J, Kavanagh D, Keller MC, Kennedy JL, Khrunin A, Kim Y, Klovins J, Knowles JA, Konte B, Kucinskis V, Kucinskiene ZA, Kuzelova-Ptackova H, Kahler AK, Laurent C, Keong JLC, Lee SH, Legge SE, Lerer B, Li MX, Li T, Liang KY, Lieberman J, Limborska S, Loughland CM, Lubinski J, Lonnqvist J, Macek M, Magnusson PKE, Maher BS, Maier W, Mallet J, Marsal S, Mattheisen M, Mattingsdal M, McCarley RW, McDonald C, McIntosh AM, Meier S, Meijer CJ, Melegh B, Melle I, Meshulam-Gately RI, Metspalu A, Michie PT, Milani L, Milanova V, Mokrab Y, Morris DW, Mors O, Murphy KC, Murray RM, Myin-Germeys I, Muller-Myhsok B, Nelis M, Nenadic I, Nertney DA, Nestadt G, Nicodemus KK, Nikitina-Zake L, Nisenbaum L, Nordin A, O'Callaghan E, O'Dushlaine C, O'Neill FA, Oh SY, Olincy A, Olsen L, Van Os J, Pantelis C, Papadimitriou GN, Papiol S, Parkhomenko E, Pato MT, Paunio T, Pejovic-Milovancevic M, Perkins DO, Pietilainen O, Pimm J, Pocklington AJ, Powell J, Price A, Pulver AE, Purcell SM, Quedsted D, Rasmussen HB, Reichenberg A, Reimers MA, Richards AL, Roffman JL, Roussos P, Ruderfer DM, Salomaa V,

- Sanders AR, Schall U, Schubert CR, Schulze TG, Schwab SG, Scolnick EM, Scott RJ, Seidman LJ, Shi JX, Sigurdsson E, Silagadze T, Silverman JM, Sim K, Slominsky P, Smoller JW, So HC, Spencer CCA, Stahl EA, Stefansson H, Steinberg S, Stogmann E, Straub RE, Strengman E, Strohmaier J, Stroup TS, Subramaniam M, Suvisaari J, Svrakic DM, Szatkiewicz JP, Soderman E, Thirumalai S, Toncheva D, Tosato S, Veijola J, Waddington J, Walsh D, Wang D, Wang Q, Webb BT, Weiser M, Wildenauer DB, Williams NM, Williams S, Witt SH, Wolen AR, Wong EHM, Wormley BK, Xi HS, Zai CC, Zheng XB, Zimprich F, Wray NR, Stefansson K, Visscher PM, Adolfsson R, Andreassen OA, Blackwood DHR, Bramon E, Buxbaum JD, Borglum AD, Cichon S, Darvasi A, Domenici E, Ehrenreich H, Esko T, Gejman PV, Gill M, Gurling H, Hultman CM, Iwata N, Jablensky AV, Jonsson EG, Kendler KS, Kirov G, Knight J, Lencz T, Levinson DF, Li QQS, Liu JJ, Malhotra AK, McCarroll SA, McQuillin A, Moran JL, Mortensen PB, Mowry BJ, Nothen MM, Ophoff RA, Owen MJ, Palotie A, Pato CN, Petryshen TL, Posthuma D, Rietschel M, Riley BP, Rujescu D, Sham PC, Sklar P, St Clair D, Weinberger DR, Wendland JR, Werge T, Daly MJ, Sullivan PF, O'Donovan MC, Consortium PG, Conso PEI, Consor WTC-C. 2014. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 511(7510):421-+.
- Rudzinski LA, Meador KJ. 2013. Epilepsy and neuropsychological comorbidities. *Continuum* 19(3 Epilepsy):682-696.
- Schmitt A, Malchow B, Hasan A, Falkai P. 2014. The impact of environmental factors in severe psychiatric disorders. *Frontiers in neuroscience* 8:19.
- Sethi NK, Emery E. 2013. Seizure with an unusual clinical semiology mimicking psychosis: a case report. *The Journal of neuropsychiatry and clinical neurosciences* 25(2):E66-67.
- Solovieff N, Cotsapas C, Lee PH, Purcell SM, Smoller JW. 2013. Pleiotropy in complex traits: challenges and strategies. *Nat Rev Genet* 14(7):483-495.
- Tellez-Zenteno JF, Patten SB, Jette N, Williams J, Wiebe S. 2007. Psychiatric comorbidity in epilepsy: a population-based analysis. *Epilepsia* 48(12):2336-2344.
- Torta R, Keller R. 1999. Behavioral, psychotic, and anxiety disorders in epilepsy: etiology, clinical features, and therapeutic implications. *Epilepsia* 40 Suppl 10:S2-20.
- Tucker GJ. 1998. Seizure disorders presenting with psychiatric symptomatology. *The Psychiatric clinics of North America* 21(3):625-635, vi.
- Vonberg FW, Bigdeli TB. 2016. Genetic Correlation Between Schizophrenia and Epilepsy. *Jama Neurol* 73(1):125-126.
- Vuilleumier P, Jallon P. 1998. [Epilepsy and psychiatric disorders: epidemiological data]. *Revue neurologique* 154(4):305-317.
- Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, Klemm A, Flicek P, Manolio T, Hindorf L, Parkinson H. 2014. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res* 42(Database issue):D1001-1006.
- Wong EHM, So HC, Li MX, Wang Q, Butler AW, Paul B, Wu HM, Hui TCK, Choi SC, So MT, Garcia-Barcelo MM, McAlonan GM, Chen EYH, Cheung EFC, Chan RCK, Purcell SM, Cherny SS, Chen RRL, Li T, Sham PC. 2014. Common Variants on Xq28 Conferring Risk of Schizophrenia in Han Chinese. *Schizophrenia Bull* 40(4):777-786.
- Yang J, Ferreira T, Morris AP, Medland SE, Madden PAF, Heath AC, Martin NG, Montgomery GW, Weedon MN, Loos RJ, Frayling TM, McCarthy MI, Hirschhorn JN, Goddard ME, Visscher PM, Trai GIA, Meta-A DGR. 2012. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nature Genetics* 44(4):369-U170.

Table 2 PRS associations from different training and target datasets

<u>Dataset Threshold</u>	SCZ PRS	EPIall PRS	EPIfoc
--------------------------	---------	------------	--------

		#SNPs	OR (99.9%CI)	P	#SNPs	OR (99.9%CI)	P	#SNPs	OR (99.9%CI)	P
HK SCZ	0.5	47610	1.62 (1.31-1.99)	2.77E-14	48486	1.06 (0.87-1.29)	0.363	44756	1.00 (0.81-1.23)	0.99
	0.05	15702	1.65 (1.33-2.04)	1.38E-14	11977	1.07 (0.88-1.31)	0.252	10701	1.03 (0.84-1.26)	0.99
	0.005	4938	1.61 (1.30-1.98)	7.13E-14	1819	0.98 (0.81-1.20)	0.787	1538	0.99 (0.81-1.20)	0.99
	Best fit	30796	1.70 (1.37-2.10)	2.26E-16	6973	1.09 (0.89-1.33)	0.161	19544	0.95 (0.77-1.17)	0.95
HK EPI	0.5	47701	1.06 (0.88-1.26)	0.310	48486	1.03 (0.86-1.24)	0.536	44761	1.08 (0.90-1.30)	0.99
	0.05	15718	1.05 (0.88-1.26)	0.347	11976	1.05 (0.88-1.26)	0.362	10704	1.10 (0.92-1.31)	0.99
	0.005	4941	1.10 (0.93-1.32)	0.065	1819	1.00 (0.83-1.19)	0.949	1536	1.10 (0.92-1.31)	0.99
	Best fit	4941	1.10 (0.93-1.32)	0.065	6973	1.10 (0.92-1.31)	0.086	13801	1.14 (0.96-1.35)	0.99

Odds ratio (OR), 99.9% confidence interval (99.9%CI) and *P*-values were calculated for standardized linear regression with adjustment for two PCs. #SNPs, number of SNPs included in PRS construction given the significance threshold.

Table 3 Heritability and genetic correlation

Dataset1	Dataset2	σ_1^2	σ_2^2	σ_{12}	R^2	<i>P</i>
SCZ	HKSCZ	0.25 (0.01)	0.23 (0.24)	0.23 (0.02)	0.39 (0.06)	4.90E-10
SCZ	HKEPI	0.25 (0.01)	0.24 (0.13)	0.04 (0.02)	0.06 (0.04)	0.116
EPIall	HKSCZ	0.08 (0.01)	0.23 (0.24)	-0.01 (0.05)	-0.06 (0.38)	0.868
EPIall	HKEPI	0.08 (0.01)	0.24 (0.13)	0.03 (0.04)	0.13 (0.23)	0.575
EPIfocal	HKSCZ	0.12 (0.01)	0.23 (0.24)	-0.05 (0.05)	-0.42 (0.76)	0.58
EPIfocal	HKEPI	0.12 (0.01)	0.24 (0.13)	-0.01 (0.05)	-0.06 (0.27)	0.818
HKSCZ	HKEPI	0.23 (0.24)	0.24 (0.13)	0.08 (0.16)	0.20 (0.34)	0.553
SCZ	EPIall	0.25 (0.01)	0.08 (0.01)	0.02 (0.01)	0.08 (0.06)	0.164
SCZ	EPIfocal	0.25 (0.01)	0.12 (0.01)	0.01 (0.01)	0.04 (0.05)	0.363
EPIall	EPIfocal	0.08 (0.01)	0.12 (0.01)	0.14 (0.02)	0.91 (0.03)	3.70E-275

σ_1^2 and σ_2^2 are heritability (in liability scale) for dataset1 and dataset2; σ_{12} and R^2 are genetic covariance and genetic correlation between dataset1 and dataset2. Numbers in bracket are standard errors for each estimate.

Table 4 Power for PRS analysis

Phenotype Pairs	Parameters ¹	Threshold			
		0.005	0.05	0.5	Best fit
SCZ, HKSCZ	0.01, 0.01, 77080, 2038, 0.43, 0.18, 0.25, 0.23, 56162, 0.88	1.000	1.000	1.000	1.000
SCZ, HKEPI	0.01, 0.003, 77080, 2138, 0.43, 0.24, 0.25, 0.04, 59821, 0.75	0.284	0.405	0.364	0.284
EPIall, HKSCZ	0.005, 0.01, 26954, 2038, 0.26, 0.18, 0.08, -0.01, 60471, 0.8	0.001	0.001	0.002	0.001
EPIall, HKEPI	0.005, 0.003, 26954, 2183, 0.26, 0.24, 0.08, 0.03, 55595, 0.95	0.072	0.066	0.045	0.071
EPIfocal, HKSCZ	0.003, 0.01, 23985, 2038, 0.17, 0.18, 0.12, -0.05, 53940, 0.86	0.011	0.032	0.058	0.042

EPIfocal, 0.003, 0.003, 23985, 2183, 0.17,
HKEPI 0.24, 0.12, -0.01, 50375, 0.95 0.004 0.003 0.003 0.003

Parameters in order for K_1 , K_2 , n_1 , n_2 , P_1 , P_2 , σ_1^2 , σ_{12} , m , and π_0 respectively.

Figure 1 Polygenic risk score association for epilepsy and schizophrenia

A) SCZ PRS on HKSCZ; B) SCZ PRS on HKEPI; C) EPIall PRS on HKSCZ; D) EPIall PRS on HKEPI; E) EPIfocal PRS on HKSCZ; F) EPIfocal PRS on HKEPI. Nagelkerke's pseudo- R^2 (Y-axis) is shown for scores derived using nine significance thresholds (X-axis). P -values from logistic regression for training PRS on target phenotype are shown above the bars, with heights representing pseudo- R^2 and color representing significance level.

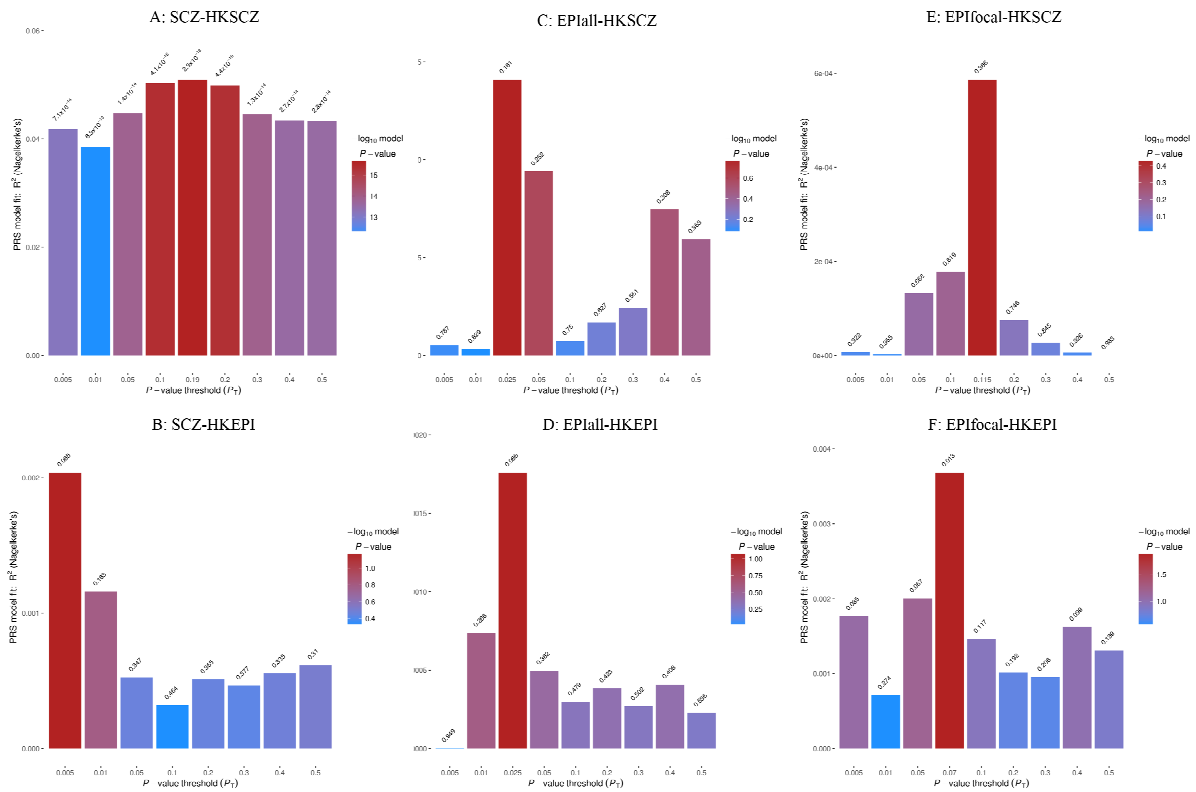
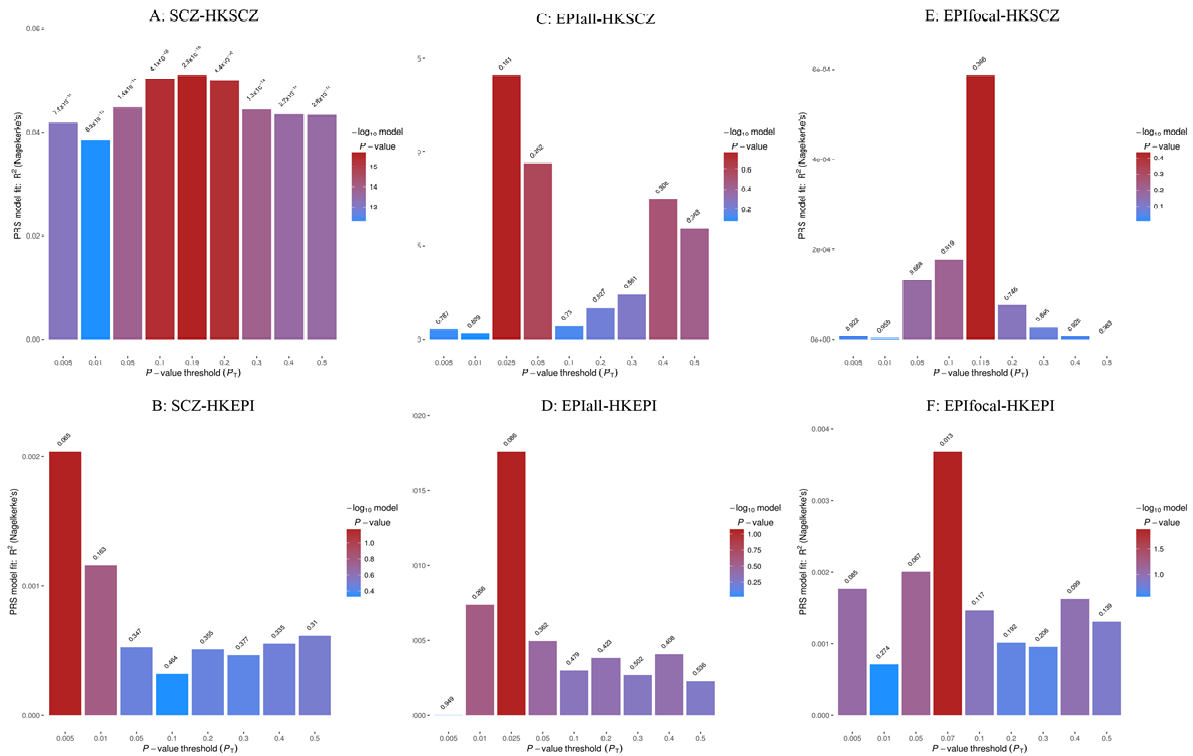


Figure 1



Gui'spaper_Figure1_revision .

Tables and Figures

Table 1 Information of data collected in this study

Dataset	Disease prevalence	Resources	Data after processing ¹
EPIall	0.005	Summary statistics from LAE meta-analysis ²	7019 cases, 19935 controls; ~1.1M SNPs
EPIfocal	0.003	Summary statistics from ILAE meta-analysis	4050 case, 17481 controls; ~1.1M SNPs
HKEPI	0.003	Raw genotypes from HK GWAS	522 cases, 1661 controls; ~800K SNPs
SCZ	0.01	Summary statistics from PGC meta-analysis ³	33356 cases, 43724 controls; ~1.1M SNPs
HKSCZ	0.01	Raw genotypes from HK GWAS	377 cases, 1661 controls; ~800K SNPs

¹ Both sample filtering and variant filtering were adopted (see method part for detail). ² Caucasian population cohorts were included for re-running meta-analysis. ³ Meta-analysis using all 49 Caucasian cohorts. ILAE, international league against epilepsy; PGC, psychiatric genomic consortium.