

ARTICLE

Open Access

The patterns of family genetic risk scores for eleven major psychiatric and substance use disorders in a Swedish national sample

Kenneth S. Kendler^{1,2}, Henrik Ohlsson³, Jan Sundquist^{3,4} and Kristina Sundquist^{3,4}

Abstract

To clarify the structure of genetic risks for 11 major psychiatric disorders, we calculated, from morbidity risks for disorders in 1st–5th degree relatives controlling for cohabitation effects, in the Swedish population born between 1932 and 1995 ($n = 5,830,014$), the family genetic risk scores (FGRS) for major depression (MD), anxiety disorders (AD), obsessive-compulsive disorder (OCD), bipolar disorder (BD), schizophrenia (SZ), bulimia (BUL), anorexia nervosa (AN), alcohol use disorder (AUD), drug use disorder (DUD), ADHD, and autism-spectrum disorder (ASD). For all affected individuals, we calculated their mean standardized FGRS for each disorder. The patterns of FGRS were quite similar for MD and AD, and for AUD and DUD, but substantially less similar for BUL and AN, BD and SZ, and ADHD and ASD. While OCD had high levels of FGRS for MD and AD, the overall FGRS profile differed considerably from MD and AD. ADHD FGRS scores were substantially elevated in AUD and DUD. FGRS scores for BD, OCD, AN, ASD, ADHD, and especially SZ were relatively disorder-specific while genetic risk for MD and AD had more generalized effects. The levels of FGRS for BMI, coronary artery disease, and educational attainment across our disorders replicated prior associations found using molecular genetic methods. All diagnostic categories examined had elevated FGRS for many disorders producing, for each condition, an informative FGRS profile. Using a novel method which approximates, from pedigree data, aggregate genetic risk, we have replicated and extended prior insights into the structure of genetic risk factors for key psychiatric illnesses.

Introduction

Several different approaches have been taken to clarify the structure of the genetic risk factors for major psychiatric disorders (including substance use disorders). Most commonly, structural equation modeling (SEM) has been applied to twin samples^{1–3} or to polygenic risk scores (PRS) generated from genome-wide association studies^{4,5}. These studies have been quite informative but have some limitations. For twin studies, sample sizes have not been insufficient to study key rarer disorders such as schizophrenia and autism spectrum disorder. For

molecular genetic studies, case–control samples have been typically assembled from a wide variety of countries, using a range of ascertainment strategies and diagnostic approaches. Furthermore, few approaches to date have utilized a different analytic approach—calculating the average magnitude of genetic risk for each proband group for a wide array of disorders.

In this report, we address some of these prior limitations by examining individuals with a lifetime diagnosis of 11 psychiatric disorders ascertained through national registries in the Swedish population born between 1932 and 1995 ($n = 5,830,014$): major depression (MD), anxiety disorders (AD), obsessive-compulsive disorder (OCD), bipolar disorder (BD), schizophrenia (SZ), bulimia (BUL), anorexia nervosa (AN), alcohol use disorder (AUD), drug use disorder (DUD), ADHD, and autism spectrum

Correspondence: Kenneth S. Kendler (Kenneth.Kendler@vcuhealth.org)

¹Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, VA, USA

²Department of Psychiatry, Virginia Commonwealth University, Richmond, VA, USA
Full list of author information is available at the end of the article

© The Author(s) 2021



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

disorder (ASD). Family genetic risk scores (FGRS) are calculated for each disorder from the weighted rates of that disorder in 1st through 5th degree relatives, controlling for cohabitation effects. We added *Family* in this term to clearly differentiate this statistic from PRS, as the FGRS derives its information not from molecular variants but from the phenotypes of a proband's *family*. This method is complimentary to PRS for the analysis of psychiatric disorders. Its current main advantage is its availability for the entire population of Sweden, coupled with excellent medical and other national registry data. Its advantage over twin and twin-family analyses conducted by structural modeling is its use of a far wider range of relatives thereby providing substantially greater information about genetic risk.

First, we examine our FGRS by disorder, comparing the patterns of the 11 FGRS scores in each of the group of probands affected with our 11 disorders one at a time. That is, a proband in this study is defined as any individual who was registered in the relevant Swedish registries with one or more of these 11 disorders. This approach allows us to examine the degree to which the genetic vulnerability of individual disorders is largely the result of an elevated FGRS for one disorder, for a few conditions or for a wide range of disorders, and the similarity of the FGRS profile, across pairs or groups of disorders.

Second, we examine our results by FGRS, ranking the scores of each specific FGRS from highest to lowest across our 11 disorders. Examined in this way, we can see the relative diagnostic specificity of each FGRS.

Third, we validate our FGRS by examining their association with three non-psychiatric phenotypes previously associated in genetic epidemiologic or molecular genetics studies with some of our psychiatric disorders: body mass index (BMI), coronary artery disease (CAD), and years of education (YOE).

Methods

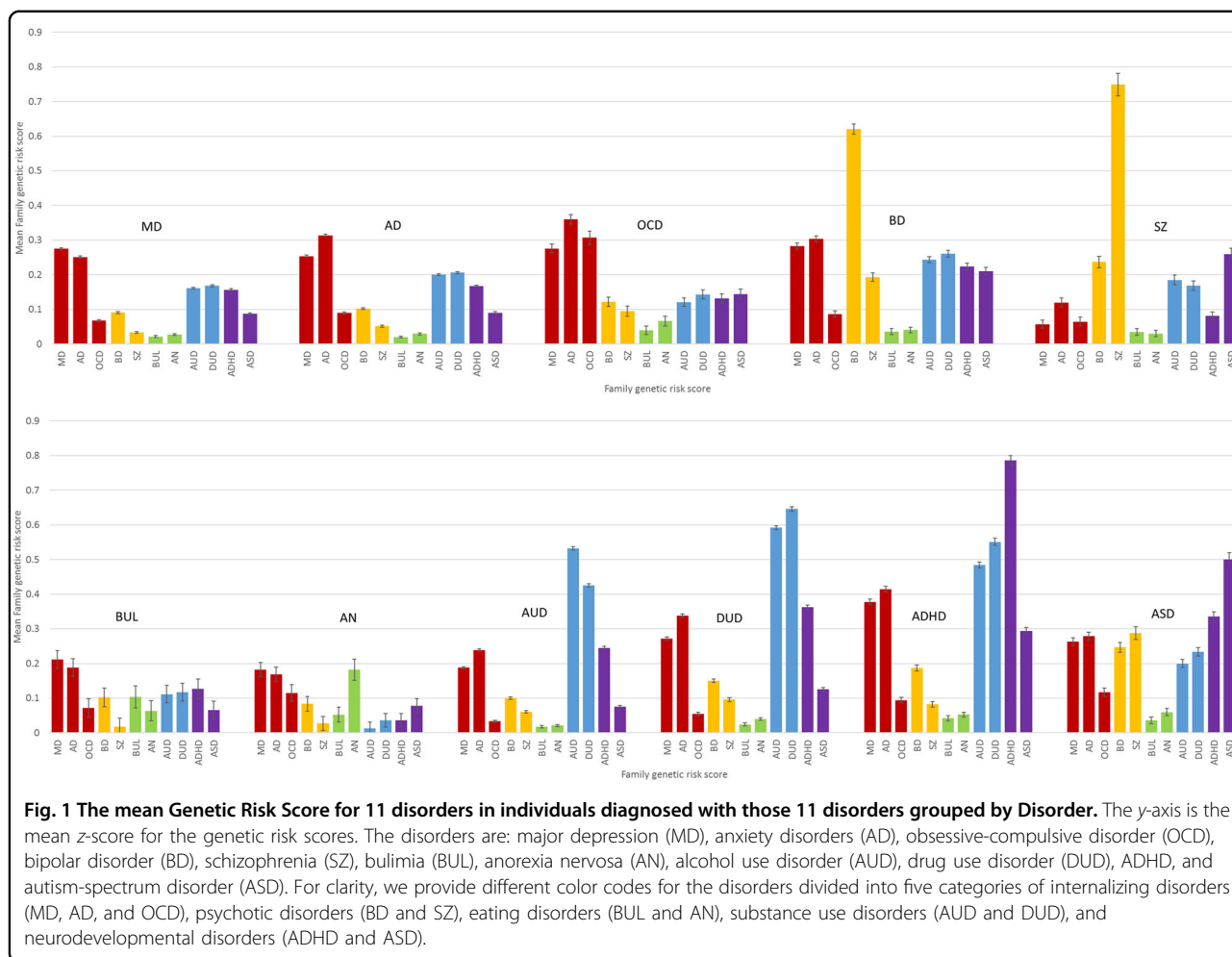
We collected information on individuals from Swedish population-based registers with national coverage linking each person's unique personal identification number which, for confidentiality, was replaced with a serial number by Statistics Sweden. This study was approved by the Regional Ethical Review Board of Lund (No. 2008/409, 2012/795, and 2016/679). Our database consisted of all individuals born in Sweden between 1932 and 1995 of

parents who themselves were born in Sweden and followed through Dec 31, 2017. We added the later requirement to ensure roughly comparable numbers of relatives for each of our probands. In the database, we included date of registration for major depression (MD), anxiety disorders (AD), obsessive-compulsive disorder (OCD), bipolar disorder (BD), schizophrenia (SZ), bulimia (BUL), anorexia nervosa (AN), alcohol use disorder (AUD), drug use

Table 1 Descriptive features of our population cohort and case samples.

		Prevalence (%)	
		All	Females Males
Total, <i>n</i>	5,830,014		
Females, <i>n</i>	2,846,911 (48.8%)		
Males, <i>n</i>	2,983,103 (51.2%)		
Mean (SD) age at follow-up	54.4 (18.1)		
		All	Females Males
Major depression	11.4	14.7	8.2
Anxiety disorders	10.6	13.7	7.7
Obsessive-compulsive disorder	0.5	0.7	0.4
Bipolar disorder	1.2	1.4	0.9
Schizophrenia	0.5	0.4	0.5
Bulimia	0.1	0.2	0.01
Anorexia	0.2	0.4	0.02
Alcohol use disorder	6.2	3.4	8.8
Drug abuse	3.5	2.5	4.5
ADHD	1.4	1.2	1.5
Autism spectrum disorder	0.7	0.6	0.8

disorder (DUD), ADHD, and autism spectrum disorder (ASD), utilizing ICD-8, 9, 10 codes from Swedish national primary care, specialist and hospital registries as well as information from Prescription and Criminal registers for AUD and DUD (see appendix Table 1 for full definitions). We also included individual genetic risk scores (FGRS) for all traits. For validation purposes, we also included FGRSs for BMI, CAD, and YOE. The FGRSs were based on selected 1st, 2nd, 3rd, 4th, and 5th degree relatives to the probands with a mean of 40.1 relatives per proband. Briefly (see appendix Fig. 1 for full details), we first calculated the morbid risk for the phenotype in our sample of relatives based on age at first registration and then we transformed the binary trait into an underlying liability distribution, with the threshold that divides the population into the two categories for the trait. Thereafter, we calculated the mean *z*-score for relatives with the trait and the mean *z*-score for individuals without. For 1st degree relatives we also multiplied the *z*-score with a factor that sought to correct for the influence of shared environmental factors separately for siblings and parent-offspring pairs. For parent-offspring pairs, this correction was implemented by comparing the resemblance, by logistic regression, for father-offspring pairs where the father sired and raised his child (that is, a father in an intact family) to the resemblance between children and their not-lived-with fathers, i.e., those who sired their offspring but never lived with or near them when they were



growing up. We have examined such not-lived-with fathers in several prior extended adoption studies^{6,7} as reflecting parent–offspring resemblance resulting only from genetic effects, analogous to the biological parent in an adoption design. For sibling pairs, we compared the resemblance in half-sibs who were versus were not reared together. Having the same genetic relationship between these siblings, this comparison also isolates the genetic from the shared environmental effects. As seen in Fig. 1 in the appendix, the correction factors—the degree of the resemblance for our individual diagnoses that was retained after discounting the effect of shared environment—varied, across diagnoses, from 0.67 to 0.99 for parent–offspring pairs and from 0.52 to 0.88 for sibling pairs.

Within each type of relative, we then had two components—the sum of the z-score and the total weighted number of relatives. These two components were weighted according to the genetic resemblance to the proband. For each proband, we summed the two components across all groups of relatives and used the quotient between the two components. Finally, to obtain the

individual FGRS, we multiplied the quotient with a shrinkage factor based on the variance of the z-score across all relatives, the variance in the mean z-score across all probands, and the number of weighted number of relatives for each proband. So that the FGRSs would be more comparable across traits and to reduce the effect of register coverage, we standardized the FGRS by year of birth into a z-score with mean = 0 and SD = 1.

For our first and second aims, we calculated the mean individual FGRSs for the 11 disorders among individuals affected, individually, with each of the 11 conditions. As a validation effort, we also examined the mean individual FGRSs for BMI, CAD, and YOIE among individuals affected with our 11 disorders. All analyses were performed using SAS 9.4⁸.

Results

Sample description

Descriptive results for our sample are seen in Table 1. Our cohort included 5,830,014 individuals with a mean (SD) age at follow-up of 54.4 (18.1). The 11 disorders had

lifetime prevalences ranging from 0.1% for BUL to 11.4% for MD. We observe the expected sex ratio with a female preponderance for MD, AD, OCD, BD, BUL, and AN, and a male preponderance for SZ, AUD, DUD, ADHD, and ASD.

Analysis by disorder

Figure 1 presents, one by one, the 11 mean standardized FGRS \pm 95% CIs for individuals affected with each of our 11 disorders. Of the many interesting patterns seen in this figure, we emphasize two. First, the general pattern of FGRS permits the division of our 11 disorders into 3 groups. For four disorders, BD, SZ, ADHD, and ASD, the FGRS scores for that disorder in affected individuals are substantially higher than that for any of the 10 other FGRS. For example, for individuals affected with SZ, the FGRS for SZ is more than three times higher than the next highest FGRS, in this case for ASD.

For four disorders, MD, AD, AUD, and DUD, their own FGRS are also highest in affected individuals but FGRS for one or more other disorders are nearly as high. For example, in individuals with MD, the mean AD FGRS score is nearly as high as the mean MD FGRS score. For subjects with DUD, the mean AUD FGRS is nearly as high as the DUD FGRS. For three disorders, OCD, BUL, and AD, the highest FGRS scores in affected individuals are *not* for the disorder themselves but instead are for MD for the two eating disorders and AD for OCD.

Second, by comparing the FGRS profile across disorders, we can observe several groups of disorders with important shared patterns of FGRS scores but often also with unique disorder-specific features. MD, AD, and OCD all share similarly elevated levels of MD and AD FGRS and the overall pattern of the other disorders is quite similar for MD and AD. OCD, by contrast, is more distinctive compared to the other two internalizing disorders, with a much higher FGRS for OCD, and also, compared with MD and AD, elevated levels of FGRS for SZ and ASD and reduced levels of FGRS for AUD and DUD. DUD and AUD have a robust reciprocal relationship where each disorder providing the second highest FGRS for the other disorder. However, DUD has higher levels for nearly every FGRS compared to AUD, particularly MD, AD, ADHD, and ASD. ADHD is a looser member of these two substance use disorders as the AUD and DUD FGRS comprise the second and third highest FGRS in ADHD subjects. SZ and ASD share some similarities in their profiles. ASD FGRS is the second highest FGRS for subjects with schizophrenia and the SZ FGRS is the third highest FGRS for individuals with ASD. Although both are eating disorders, the FGRS profiles for BUL and AN share some similarities but many differences. The highest FGRS for both disorders are for MD and AD. BUL has much higher levels of AUD, DUD, and ADHD FGRS than does AN, while AN has higher OCD

FGRS. Furthermore, the FGRS for BUL is quite low in subjects with AN. Although both are often considered “psychotic disorders”, the FGRS profiles for BD and SZ are quite different. FGRS scores for MD, AD, and ADHD are considerably higher in the BD versus SZ subjects and the cross-loading of their own FGRS scores (that is levels of BD FGRS in SZ subjects and SZ FGRS in BD subjects) is much more modest than are seen with the MD–AD and AUD–DUD pairs.

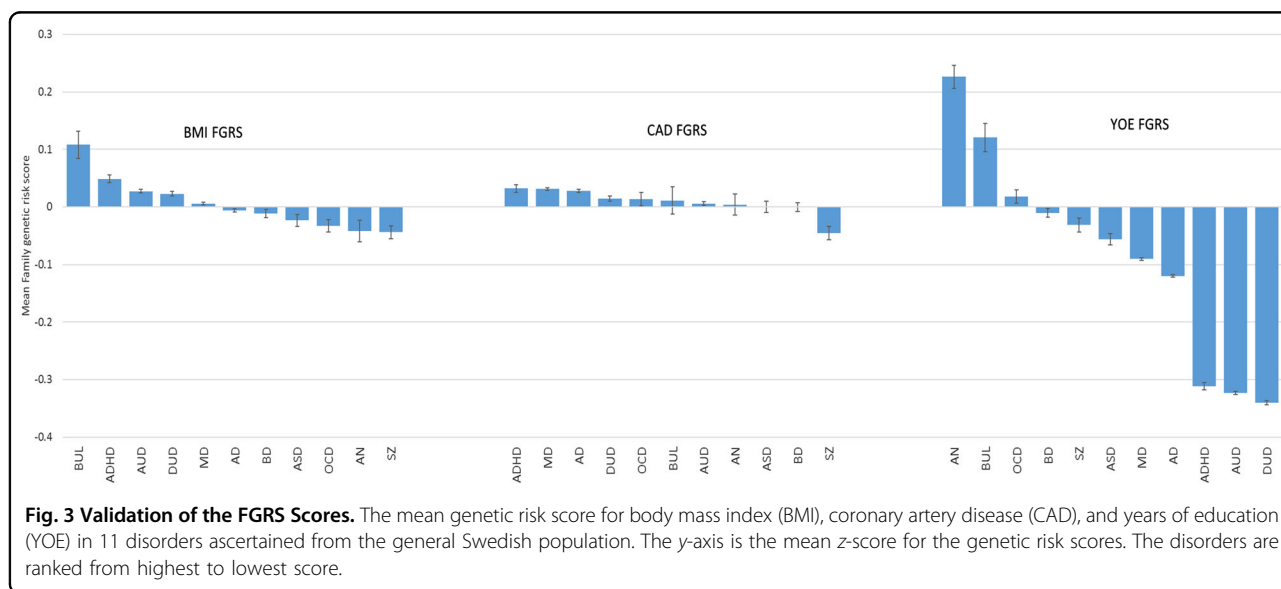
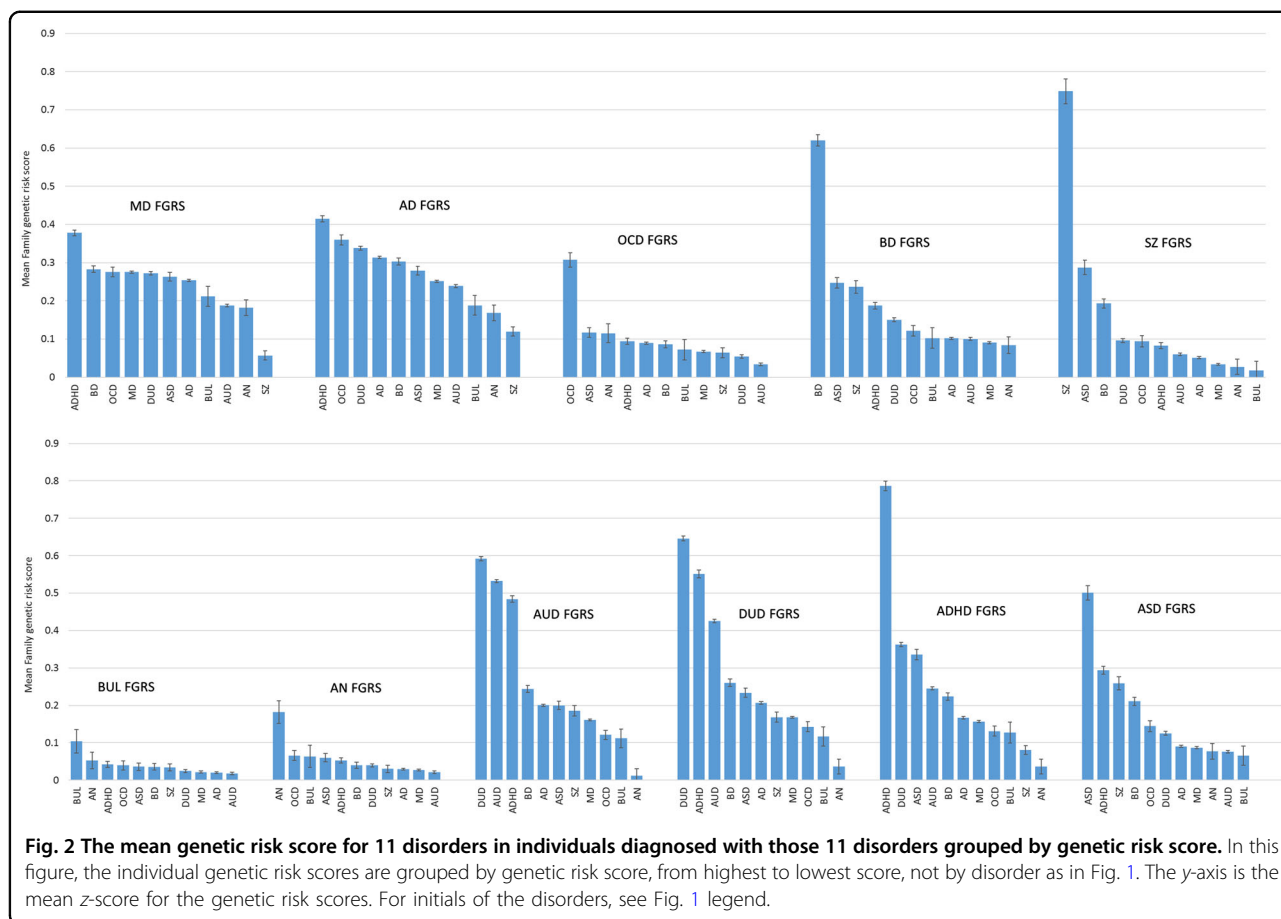
Analysis by genetic risk score

Figure 2 presents the same data but arranged by FGRS rather than by disorder. For each FGRS, the disorders are ranked by FGRS from highest to lowest. We here emphasize two broad points from these results. First, we can examine the pattern of the specificity of each FGRS by two criteria: (i) is the FGRS scores highest in the disorder from which the FGRS is created? If yes, (ii) is there a substantial difference between the mean FGRS for that disorder and for the disorder with the second highest FGRS which we operationalize as a minimum 2:1 ratio? Eight of the 11 disorders meet the first criterion of specificity, all but MD, AD, and AUD. MD failed because subjects with ADHD have a slightly higher mean FGRS for MD than those with MD. AD fails because individuals with both ADHD and OCD have an AD FGRS slightly higher than do AD subjects. The next criterion was met by six of the eight disorders: OCD, BD, SZ, BUL, AN, and ADHD. DUD did not meet these conditions because of high levels of DUD FGRS in subjects with both ADHD and AUD. The ASD FGRS did not meet this criterion because of the relatively high levels of ASD FGRS in subjects with ADHD.

Second, the findings displayed by FGRS in Fig. 2 permit other observations of patterns of resemblance across disorders based on the relative levels of their particular FGRS scores. For example, patients with ASD had the second highest FGRS scores for BD and SZ and the third highest FGRS score for ADHD. Patients with ADHD had the second highest FGRS scores for ASD and DUD and the 3rd highest score for AUD. Aside from SZ itself, only two other disorders had appreciably increased risk for SZ FGRS, ASD and BD. Individuals with AN had the lowest FGRS scores for AUD, DUD, and ADHD and the second lowest for SZ and AD.

Validation of FGRS scores

Given the novelty of this method, our final analyses attempt to validate our FGRS method by examining, across our 11 disorders, the FGRS scores for BMI, CAD, and YOE (Fig. 3). For BMI, the effects were modest with the highest and lowest FGRS scores seen for BUL and AN, respectively. For CAD, the effects were even smaller, with AN and OCD having the highest mean risk and SZ the lowest. The effects were much larger for YOE where high



YOE FGRS were seen, in order, for AN, BUL, OCD, BD, and ASD. Lower than average scores were seen, from largest to least reduction from the mean, for DUD, AUD, ADHD, AD, MD, and SZ.

Discussion

Of the many points of possible interest in these results, we review four. *First*, a number of prior approaches, using latent variables applied to genetic correlations, have

examined the structure of genetic risks for major psychiatric disorders^{1–5}. Using a different method, our results also provide nuanced evidence for the inter-relationships of the genetic risks for different disorders. The two most striking examples are the close relationships revealed between MD and AD, and between AUD and DUD.

For MD and AD, consistent with prior evidence for a close genetic association between these two “internalizing disorders”^{1,3,5}, the relationship is evidenced both by a reciprocal relationship between their FGRS scores (AD FGRS is the second highest score for MD patients and MD FGRS is the second highest FGRS for AD patients), and by a quite similar profile of their other FGRS scores. It is instructive to compare their two profiles with that of OCD which, when included in phenotypic analyses, typically aligns in factor analyses with other internalizing disorders^{9,10} but in a recent molecular genetic multivariate analysis it was shown to cluster instead with AN and Tourette’s syndrome⁴. With AD and MD, OCD shares high levels of AD and MD FGRS, but the pattern of other FGRS are quite different including, consistent with prior studies^{4,11}, elevated levels of genetic risk for AN.

Congruent with prior evidence for shared genetic underpinnings for externalizing disorders^{1,3,5}, the strong relationship between AUD and DUD is marked by sharing the same two highest FGRS for AUD and DUD. But differences in the FGRS profiles of these two drug use disorders are also evident. In particular, DUD, compared to AUD, has considerably higher levels of FGRS for MD, AD, BD, SZ, and ADHD.

However, our findings did not confirm close genetic relationships between two other pairs of disorders expected to be closely genetically related: AN and BUL, and BD and SZ. Family and twin studies show substantial sharing of genetic risk between AN and BUL^{12,13}. However, the FGRS loadings for AN in BUL cases, and BUL in AN cases, were rather modest. Their genetic profiles also differed in other ways. For example, consistent with prior studies, BUL had a substantially higher FGRS scores for the externalizing disorders of AUD, DUD, and ADHD than did AN^{14,15}.

Our results are also not consistent with recent reports of high genetic correlations between SZ and BD^{16,17}, but are with earlier family studies^{18–20} showing quite modest co-aggregation. While the FGRS score of BD is the second highest observed for SZ, it is much weaker, and the association is not reciprocal as the SZ FGRS is the eighth strongest FGRS for BD. Furthermore, the profile for other FGRS scores differs meaningful across the two disorders.

Our results do provide support for a modest reciprocal genetic relationship between SZ and ASD reported elsewhere^{17,21} as the ASD FGRS is the second strongest for SZ and the SZ FGRS is the third strongest for ASD. Also, consistent with prior molecular genetics’ findings^{4,5}, our

two neurodevelopmental disorders have a reciprocal genetic inter-relationship as the ADHD FGRS is the second strongest for ASD, and ASD FGRS the third strongest for ADHD.

Second, the findings as presented in Fig. 2 help clarify the degree of disorder specificity of our 11 FGRS. This can be best illustrated by comparing the pattern of results for the FGRS for SZ and MD. The profile for SZ is dominated by a single substantial FGRS for SZ itself nearly three times greater higher than the next strongest FGRS, which is for ASD. By contrast, the curve for the MD FGRS is flatter with five disorders having MD FGRS scores only modestly lower than that seen for MD itself. A second sign of the relative non-specificity of risk to MD is that the MD FGRS is not highest for MD itself. The other FGRS with a similar pattern of non-specificity is AD. Our results suggest that these two classical internalizing disorders²² have a genetic substrate somewhat different from the other disorders examined. They substantially predispose to two disorders that are common, often disabling and are leading causes of world-wide disability²³. However, these genetic risk factors also make important contributions to a wider range of disorders including BD, ADHD, and substance use and eating disorders.

Third, given the novelty of our FGRS methods, we sought to validate them in three non-psychiatric phenotypes to see if we could replicate previously observed genetic associations. We examined BMI because of intriguing evidence that genetic risk for AN was associated with constitutional thinness^{24,25}. We replicated that finding showing a stronger inverse association of AN FGRS with BMI than any of the other disorders. Consistent with prior reports, we also found a negative association between our FGRS for BMI and FGRS for SZ²⁶ and BD²⁵, and a positive association between the BMI FGRS and ADHD²⁵. We also showed, for the first time to our knowledge, a positive correlation between the FGRS for BMI and BN, consistent with the finding that a BMI polygene risk score predicted binge eating and purging in a general population sample²⁷.

Prior genetic epidemiological and molecular genetic data have found an association between risk for MD and risk for CAD^{28,29}. We also replicated those results. Finally, previous studies have found a positive genetic correlation between the genetic predisposition to high educational attainment and risk for ASD, BD, and AN^{24,30,31} and a negative genetic correlation with AUD³² and DUD³³. We could reproduce all these associations. The ability of FGRS to consistently replicate a range of previously reported genetic correlations supports the validity of this method.

Fourth, perhaps the greatest contribution of these analyses to insightful prior work on the genetic relationships between disorders has been the value of examining, for

our patient cohorts, not only their genetic risk for the disorder from which they suffer, but also the pattern of their genetic risks for multiple other disorders. Our results suggest that examining such genetic profiles, we can gain a richer understanding of the genetic substrate for our disorders and the genetic relationship across disorders.

Limitations

These findings should be viewed in the context of eight potential methodological limitations. First, the validity of the FGRS score is dependent on the quality of the available diagnoses in the Swedish national registries which has been well demonstrated for SZ, BD, and OCD^{34–37}. The validity of MD diagnoses is supported by its prevalence, sex ratio, sibling and twin correlations, and associated psychosocial risk factors^{38,39}. Genetic epidemiological findings for AUD, DUD, and eating disorders in Sweden have been similar to those found in other samples^{7,13,40–42}. We are unaware of attempts to validate Swedish diagnoses for ADHD, AN, ASD, or BUL.

Second, ASD had a usually low male-to-female ratio of 1.33:1. We therefore examined this ratio as a function of age at first registration (appendix Fig. 2). Of note, in those first diagnosed under age 10, the ratio was in the expected range: > 3:1.

Third, we did not attempt to account formally for assortative mating in our analyses. However, the impact of any such spousal concordance is accounted for in our analyses in that an individual's FGRS is calculated from both maternal and paternal relatives.

Fourth, we did not correct for the fact that many individuals in the Swedish population would appear multiple times in our risk scores. They could, for example, be a proband if affected with one of our 11 disorders, and also be a sibling, parent, uncle or cousin if they had affected relatives. We do not expect this approach to produce biases in our FGRS scores, but our confidence intervals are likely to be modestly underestimated.

Fifth, our diagnoses required individuals to present for medical treatment, have a criminal contact related to AUD or DUD, abused prescription drugs or took specific pharmacological treatments for AUD. We are therefore likely to miss some mildly affected individuals and cannot rule out correlated treatment seeking in relatives as a potential confounder.

Sixth, in comparing results across a wide range of disorders, the question of diagnostic hierarchies arises. We took a minimalist approach, only utilizing a single DSM-5 hierarchy⁴³: individuals with both MD and BD are coded as BD. The other diagnostic pairing of concern was SZ and BD, where 14.0% of those with SZ also had a BD diagnosis and 5.6% with BD also had a SZ diagnosis. For those diagnoses, we developed a hierarchy and included

that with overall changes in results (see appendix Table 2 and Fig. 3).

Seventh, our control for cohabitation effects, detailed in appendix Fig. 1 step 4, are based on large samples of fathers from intact families versus not-lived-with fathers⁶ and half-sibs reared together and apart⁴⁴. Across disorders, shared environmental effects contributed a mean of $19.8 \pm 5.8\%$ to parent–offspring and $23.7 \pm 3.4\%$ to sibling–sibling resemblance, for which we corrected in our analyses. These effects are only approximate, and we did not control for environmental influences on more distant relatives. The aggregate impact of these correction on our FGRS, as seen in appendix Table 3, was quite modest.

Finally, to explore the stability of our FGRS, we examined cohort and geographical differences in all 11 of our FGRS scores in appendix Fig. 4a, b. Reassuringly, we found only modest effects of time and space on our FGRS.

Conclusions

Prior efforts to investigate the structure of genetic risk factors for psychiatric disorders have largely relied on structural equation analyses of twin samples or polygenic risk scores derived from large, often diverse, case–control cohorts. Our approach is different and hopefully complementary, utilizing rates of psychiatric illness in extended pedigrees of 5.8 million Swedes. An analysis of the resulting patterns showed that all diagnostic categories had elevated FGRS for multiple disorders. These FGRS profiles provided important insights into the structure of the genetic substrates of our major psychiatric disorders in many, but not all instances, replicating patterns of genetic sharing found using other methods. We can be particularly confident in these replicated findings given the persuasive argument that obtaining similar results using different methods (i.e., “triangulation”), is of greater value than replicating findings using the same methods⁴⁵.

Acknowledgements

This project was supported by grants AA023534 and DA030005 from the National Institutes of Health, the Swedish Research Council as well as Avtal om Läkarutbildning och Forskning (ALF) funding from Region Skåne.

Author details

¹Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, VA, USA. ²Department of Psychiatry, Virginia Commonwealth University, Richmond, VA, USA. ³Center for Primary Health Care Research, Lund University, Malmö, Sweden. ⁴Department of Family Medicine and Community Health, Department of Population Health Science and Policy, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Conflict of interest

The authors declare no competing interests.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41398-021-01454-z>.

Received: 23 April 2021 Revised: 1 May 2021 Accepted: 12 May 2021

Published online: 27 May 2021

References

- Kendler, K. S. et al. The structure of genetic and environmental risk factors for syndromal and subsyndromal common DSM-IV Axis I and all Axis II disorders. *Am. J. Psychiatry* **168**, 29–39 (2011).
- Waldman, I. D., Poore, H. E., van Hulle, C., Rathouz, P. J. & Lahey, B. B. External validity of a hierarchical dimensional model of child and adolescent psychopathology: tests using confirmatory factor analyses and multivariate behavior genetic analyses. *J. Abnorm. Psychol.* **125**, 1053 (2016).
- Cosgrove, V. E. et al. Structure and etiology of co-occurring internalizing and externalizing disorders in adolescents. *J. Abnorm. Child Psychol.* **39**, 109–123 (2011).
- Cross-Disorder Group of the Psychiatric Genomics Consortium. Genomic relationships, novel loci, and pleiotropic mechanisms across eight psychiatric disorders. *Cell* **179**, 1469–1482.e11 (2019).
- Waldman, I. D., Poore, H. E., Luningham, J. M. & Yang, J. Testing structural models of psychopathology at the genomic level. *World Psychiatry* **19**, 350–359 (2020).
- Kendler, K. S., Ohlsson, H., Sundquist, K. & Sundquist, J. Sources of parent–offspring resemblance for major depression in a national Swedish extended adoption study. *JAMA Psychiatry* **75**, 194–200 (2018).
- Kendler, K. S. et al. An extended Swedish national adoption study of alcohol use disorder. *JAMA Psychiatry* **72**, 211–218 (2015).
- SAS Institute. *I. SAS/STAT® Online Documentation, Version 9.4*. (SAS Institute, Inc., Cary, NC, 2012).
- Krueger, R. F., Caspi, A., Moffitt, T. E. & Silva, P. A. The structure and stability of common mental disorders (DSM-III-R): a longitudinal-epidemiological study. *J. Abnorm. Psychol.* **107**, 216–227 (1998).
- Miller, M. W., Fogler, J. M., Wolf, E. J., Kaloupek, D. G. & Keane, T. M. The internalizing and externalizing structure of psychiatric comorbidity in combat veterans. *J. Trauma. Stress* **21**, 58–65 (2008).
- Cederlöf, M. et al. Etiological overlap between obsessive-compulsive disorder and anorexia nervosa: a longitudinal cohort, multigenerational family and twin study. *World Psychiatry* **14**, 333–338 (2015).
- Strober, M., Freeman, R., Lampert, C., Diamond, J. & Kaye, W. Controlled family study of anorexia nervosa and bulimia nervosa: evidence of shared liability and transmission of partial syndromes. *Am. J. Psychiatry* **157**, 393–401 (2000).
- Bulik, C. M. et al. Understanding the relation between anorexia nervosa and bulimia nervosa in a Swedish national twin sample. *Biol. Psychiatry* **67**, 71–77 (2010).
- Bulik, C. M. Drug and alcohol abuse by bulimic women and their families. *Am. J. Psychiatry* **144**, 1604–1606 (1987).
- Lilenfeld, L. R. et al. A controlled family study of anorexia nervosa and bulimia nervosa: psychiatric disorders in first-degree relatives and effects of proband comorbidity. *Arch. Gen. Psychiatry* **55**, 603–610 (1998).
- Cross-Disorder Group of the Psychiatric Genomics Consortium. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat. Genet.* **45**, 984–994 (2013).
- Richardson, T. G., Harrison, S., Hemani, G. & Smith, G. D. An atlas of polygenic risk score associations to highlight putative causal relationships across the human genome. *eLife* **8**, e43657 (2019).
- Kendler, K. S. et al. The Roscommon Family Study. IV. Affective illness, anxiety disorders, and alcoholism in relatives. *Arch. Gen. Psychiatry* **50**, 952–960 (1993).
- Kendler, K. S., Gruenberg, A. M. & Tsuang, M. T. Psychiatric illness in first-degree relatives of schizophrenic and surgical control patients. A family study using DSM-III criteria. *Arch. Gen. Psychiatry* **42**, 770–779 (1985).
- Maier, W. et al. Continuity and discontinuity of affective disorders and schizophrenia. Results of a controlled family study. *Arch. Gen. Psychiatry* **50**, 871–883 (1993).
- Smoller, J. W. et al. Psychiatric genetics and the structure of psychopathology. *Mol. Psychiatry* **24**, 409–420 (2019).
- Kendler, K. S. & Myers, J. The boundaries of the internalizing and externalizing genetic spectra in men and women. *Psychol. Med.* **44**, 647–655 (2014).
- Vos, T. et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* **380**, 2163–2196 (2012).
- Duncan, L. et al. Significant locus and metabolic genetic correlations revealed in genome-wide association study of anorexia nervosa. *Am. J. Psychiatry* **174**, 850–858 (2017).
- Bulik-Sullivan, B. et al. An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).
- Solmi, F., Mascarell, M. C., Zammit, S., Kirkbride, J. B. & Lewis, G. Polygenic risk for schizophrenia, disordered eating behaviours and body mass index in adolescents. *Br. J. Psychiatry* **215**, 428–433 (2019).
- Abdulkadir, M. et al. Polygenic score for body mass index is associated with disordered eating in a general population cohort. *J. Clin. Med.* **9**, 1187 (2020).
- Consortium, T. B. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat. Genet.* **50**, 668–681 (2018).
- Kendler, K. S., Fiske, A., Gardner, C. O. & Gatz, M. Delineation of two genetic pathways to major depression. *Biol. Psychiatry* **65**, 808–811 (2009).
- Grove, J. et al. Identification of common genetic risk variants for autism spectrum disorder. *Nat. Genet.* **51**, 431–444 (2019).
- Stahl, E. A. et al. Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat. Genet.* **51**, 793–803 (2019).
- Zhou, H. et al. Genome-wide meta-analysis of problematic alcohol use in 435,563 individuals yields insights into biology and relationships with other traits. *Nat. Neurosci.* **23**, 801–818 (2020).
- Zhou, H. et al. Association of OPRM1 functional coding variant with opioid use disorder: a genome-wide association study. *JAMA Psychiatry* **77**, 1072–1080 (2020).
- Lichtenstein, P. et al. Recurrence risks for schizophrenia in a Swedish national cohort. *Psychological Med.* **36**, 1417–1425 (2006).
- Sellgren, C., Landén, M., Lichtenstein, P., Hultman, C. M. & Langstrom, N. Validity of bipolar disorder hospital discharge diagnoses: file review and multiple register linkage in Sweden. *Acta Psychiatr. Scand.* **124**, 447–453 (2011).
- Ekholm, B. et al. Evaluation of diagnostic procedures in Swedish patients with schizophrenia and related psychoses. *Nord. J. Psychiatry* **59**, 457–464 (2005).
- Rück, C. et al. Validity and reliability of chronic tic disorder and obsessive-compulsive disorder diagnoses in the Swedish National Patient Register. *BMJ Open* **5**, e007520 (2015).
- Kendler, K. S., Ohlsson, H., Lichtenstein, P., Sundquist, J. & Sundquist, K. The genetic epidemiology of treated major depression in Sweden. *Am. J. Psychiatry* **175**, 1137–1144 (2018).
- Sundquist, J., Ohlsson, H., Sundquist, K. & Kendler, K. S. Common adult psychiatric disorders in Swedish primary care (where most mental health patients are treated). *BMC Psychiatry* **17**, 235 (2017).
- Kendler, K. S. et al. A national Swedish twin-sibling study of alcohol use disorders. *Twin Res. Hum. Genet.* **19**, 430–437 (2016).
- Kendler, K. S. et al. Genetic and familial environmental influences on the risk for drug abuse: a national Swedish adoption study. *Arch. Gen. Psychiatry* **69**, 690–697 (2012).
- Kendler, K. S., Maes, H. H., Sundquist, K., Ohlsson, H. & Sundquist, J. Genetic and family and community environmental effects on drug abuse in adolescence: a Swedish national twin and sibling study. *Am. J. Psychiatry* **171**, 209–217 (2013).
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th edn, DSM-5 (American Psychiatric Association, Washington, DC, 2013).
- Maes, H. H. et al. Modeling etiology of smoking during pregnancy in Swedish twins, full- & half-siblings, reared together and apart. *Nicotine Tob. Res.* **22**, 1736–1743 (2020).
- Munafò, M. R. & Davey-Smith, G. Robust research needs many lines of evidence. *Nature* **553**, 399–401 (2018).