

Comparing Copy Number Variations in a Danish Case Cohort of Individuals With Psychiatric Disorders

Xabier Calle Sánchez, MS; Dorte Helenius, PhD; Jonas Bybjerg-Grauholm, PhD; Carsten Pedersen, PhD; David M. Hougaard, MD; Anders D. Børglum, PhD; Merete Nordentoft, PhD; Ole Mors, MD; Preben B. Mortensen, DrMedSc; Daniel H. Geschwind, MD; Simone Montalbano, MS; Armin Raznahan, MD; Wesley K. Thompson, PhD; Andrés Ingason, PhD; Thomas Werge, PhD

[+ Supplemental content](#)

IMPORTANCE Although the association between several recurrent genomic copy number variants (CNVs) and mental disorders has been studied for more than a decade, unbiased, population-based estimates of the prevalence, disease risks and trajectories, fertility, and mortality to contrast chromosomal abnormalities and advance precision health care are lacking.

OBJECTIVE To generate unbiased, population-based estimates of prevalence, disease risks and trajectories, fertility, and mortality of CNVs implicated in neuropsychiatric disorders.

DESIGN, SETTING, AND PARTICIPANTS In a population-based case-cohort study, using the Lundbeck Foundation Initiative for Integrative Psychiatric Research (iPSYCH) 2012 database, individuals born between May 1, 1981, and December 31, 2005, and followed up until December 31, 2012, were analyzed. All individuals (n = 57 377) with attention-deficit/hyperactivity disorder (ADHD), major depressive disorder (MDD), schizophrenia (SCZ), autism spectrum disorder (ASD), or bipolar disorder (BPD) were included, as well as 30 000 individuals randomly drawn from the database. Data analysis was conducted from July 1, 2017, to September 7, 2021.

EXPOSURES Copy number variants at 6 genomic loci (1q21.1, 15q11.2, 15q13.3, 16p11.2, 17p12, and 17q12).

MAIN OUTCOMES AND MEASURES Population-unbiased hazard ratio (HR) and survival estimates of CNV associations with the 5 ascertained psychiatric disorders, epilepsy, intellectual disability, selected somatic disorders, fertility, and mortality.

RESULTS Participants' age ranged from 1 to 32 years (mean, 12.0 [IQR, 6.9] years) during follow-up, and 38 662 were male (52.3%). Copy number variants broadly associated with an increased risk of autism spectrum disorder and ADHD, whereas risk estimates of SCZ for most CNVs were lower than previously reported. Comparison with previous studies suggests that the lower risk estimates are associated with a higher CNV prevalence in the general population than in control samples of most case-control studies. Significant risk of major depressive disorder (HR, 5.8; 95% CI, 1.5-22.2) and sex-specific risk of bipolar disorder (HR, 17; 95% CI, 1.5-189.3, in men only) were noted for the 1q21.1 deletion. Although CNVs at 1q21.1 and 15q13.3 were associated with increased risk across most diagnoses, the 17p12 deletion consistently conferred less risk of psychiatric disorders (HR 0.4-0.8), although none of the estimates differed significantly from the general population. Trajectory analyses noted that, although diagnostic risk profiles differed across loci, they were similar for deletions and duplications within each locus. Sex-stratified analyses suggest that pathogenicity of many CNVs may be modulated by sex.

CONCLUSIONS AND RELEVANCE The findings of this study suggest that the iPSYCH population case cohort reveals broad disease risk for some of the studied CNVs and narrower risk for others, in addition to sex differential liability. This finding on genomic risk variants at the level of a population may be important for health care planning and clinical decision making, and thus the advancement of precision health care.

Author Affiliations: Author affiliations are listed at the end of this article.

Corresponding Authors: Andrés Ingason, PhD, Institute of Biological Psychiatry, Mental Health Services, Copenhagen University Hospital, Copenhagen, Denmark (andres.ingason@regionh.dk); Thomas Werge, PhD, Institute of Biological Psychiatry, Mental Health Services, Copenhagen University Hospital, Copenhagen, Denmark, Mental Health Services, 2, Boserupvej, Roskilde 4000, Denmark (thomas.werge@regionh.dk).

JAMA Psychiatry. 2022;79(1):59-69. doi:10.1001/jamapsychiatry.2021.3392
Published online November 24, 2021. Corrected on January 5, 2022.

Over the past decades, advances in large-scale genotyping and sequencing have led to the identification of pathogenic genomic variations, such as single nucleotide variants, indels, and copy number variation (CNV), greatly expanding our knowledge of genomic disorders and their links to neuropsychiatric phenotypes.^{1,2} In particular, recurrent de novo CNVs constitute an important type of genetic variation that confers high risk for severe neurodevelopmental disorders, such as intellectual disability (ID), autism spectrum disorder (ASD), attention-deficit/hyperactivity disorder (ADHD), and later-onset psychotic disorders, as well as epilepsy and premature death.³

Although CNVs at several distinct loci have been definitively associated with neuropsychiatric disorders in large, case-control studies of specific mental diagnoses,¹ accurate population prevalence and relative risk estimates are lacking in most cases, as few reports have been based on epidemiological study designs examining genomic risk variants in a population. Accuracy in estimation of penetrance across multiple pathogenic CNVs is crucial for understanding the architecture of genetic risk in psychiatry and for our capacity to use an individual's CNV status to advance precision psychiatry and clinical diagnostics. To our knowledge, there have been no comparative population-based studies of prevalence and penetrance across the major recurrent CNVs that impart risk for psychopathologic disorders.

The need to address this large gap in knowledge is demonstrated by studies of velocardiofacial syndrome (deletion 22q11.2), which has been known for a quarter of a century and is among the best characterized of human mutations.⁴ Deletions at 22q11.2 were thought to confer a more than 20-fold increase in the risk for schizophrenia (SCZ) based on classical studies in clinically ascertained cohorts.⁴⁻⁸ However, this penetrance estimate has been substantially revised to an approximately 2-fold increase by a population-based study of CNVs at 22q11.2 in the Danish Lundbeck Foundation Initiative for Integrative Psychiatric Research (iPSYCH) case cohort.⁹ This observation, which to our knowledge is from the first population-based study of psychiatric CNV penetrance using nationwide diagnostic registry data, suggests that clinical cohorts and case-control samples are biased toward carriers who are severely affected by 22q11.2 deletion and healthy controls, and that clinically unaffected carriers are much more frequent in the general population than anticipated. It is not known whether similarly major revisions of penetrance would be seen for population-based analysis of other recurrent CNVs in psychiatry.

Herein, we used the Danish population-based iPSYCH2012 case cohort^{9,10} to estimate the true population prevalence, early mortality, and trajectory of disease risk for recurrent CNVs at 6 of the most commonly studied genomic loci (1q21.1, 15q11.2, 15q13.3, 16p11.2, 17p12, and 17q12),¹¹⁻¹⁶ as well as to directly compare their penetrance and expressivity (eTable 1 in the Supplement provides hg19 coordinates and number of probes at each locus). This iPSYCH case cohort includes all individuals born in Denmark from May 1, 1981, to December 31, 2005 (n = 57 377) diagnosed with ADHD, major depressive disorder (MDD), SCZ, ASD, or bipolar disorder (BPD), as well

Key Points

Question What are the population-based prevalence and risk of psychiatric disorders associated with pathogenic copy number variations (CNVs) and how do they compare?

Findings In a cohort study including 86 189 individuals, increased CNV-associated risk of autism, attention-deficit/hyperactivity disorder, schizophrenia, and major depressive disorder, as well as bipolar disorder in men for deletion at 1q21.1, was observed. Population-based penetrance estimates were generally lower than those from prior studies; time-dependent analyses identified variegated disease trajectories across genomic loci, whereas deletions and duplications within each locus had similar trajectory patterns.

Meaning The findings of this study suggest that population-based analysis substantially revises prevalence and penetrance estimates for pathogenic CNVs; precision health care needs to be tailored to the specific CNV, and to the age and gender of the affected individual.

as 30 000 population controls randomly drawn from the same birth cohort, corresponding to roughly 2% of the entire Danish population born in 1981-2005.¹⁰

Methods

This study was based on the iPSYCH case cohort, which has been previously described in detail.^{9,10} The use of all data in our study follows the standards of the Danish Scientific Ethics Committee, the Danish Health Data Authority, the Danish Data Protection Agency, and the Danish Neonatal Screening Biobank Steering Committee. For this study, the Danish Scientific Ethics Committee has, in accordance with Danish legislation, waived the need for informed consent in biomedical research based on existing biobanks. Data access was via secure portals in accordance with Danish data protection guidelines set by the Danish Data Protection Agency, the Danish Health Data Authority, and Statistics Denmark.

The present study is based on the iPSYCH2012 case cohort of 86 189 individuals from the 1 472 762 singletons born between May 1, 1981, and December 31, 2005, in Denmark who were residents of Denmark at age 1 year and had a mother registered in the Danish Civil Registration System^{9,10}; data on race and ethnicity are not recorded or used anywhere in Danish society. The case cohort is made up of 2 components: cases and population comparison cohort. Cases included all individuals (n = 57 377) who have been clinically diagnosed with MDD (*International Statistical Classification of Diseases, Tenth Edition [ICD-10]* code F32-33), ASD (*ICD-10* code F84), BPD (*ICD-10* code F30-31), SCZ (*ICD-10* code F20), or ADHD (*ICD-10* code F90) according to inpatient and outpatient discharge diagnoses from all Danish hospitals until December 31, 2012, obtained from the Danish Psychiatric Central Research Register.¹⁷ The total number of cases is less than the sum of the subtotal for each diagnosis because some individuals have more than 1 diagnosis, and the total number of samples in the study is smaller than the sum of cases and

cohort because the cohort is a random sample of the population and therefore includes a small number of cases (eFigure 6 in the Supplement). Details on DNA extraction and CNV calling are given in the eMethods and eFigure 1 in the Supplement.

Statistical Analysis

Primary analyses were weighted Cox proportional hazards (CPH) regression models, with age at first hospital admission as an inpatient or outpatient for each of the 5 psychiatric diagnoses as the outcomes. If individuals did not have a relevant hospital contact by December 31, 2012, or died, immigrated, or were lost to follow-up, they were censored at the appropriate date. For those with epilepsy and ID, we derived weights and analyzed data in the same way as the ascertained diagnoses, because epilepsy and ID are frequently comorbid brain disorders, especially with ASD. The presence of a deletion or duplication was included as an independent variable, coded as *no CNV* (ie, both chromosomes intact) as the reference category. All CPH regression models included sex at birth as a stratification variable. Inverse probability of sampling weights was computed as described by Barlow et al¹⁸ and used in analyses to obtain population-unbiased estimates of relative and absolute risk. A robust estimator of variance was used for computing SEs of regression coefficients. Wald statistics using the robust SEs were applied to test for significance of hazard ratios (HRs) for deletions and duplications. Sex-specific analyses were done following the same weighted CPH regression model but each sex was analyzed separately.

The multistate model was fitted using the same procedure as in the primary analyses and defined outcomes as the age at onset for each of the possible transition states: ADHD alone, ASD alone, ADHD and ASD, ADHD and ID, and ASD and ID, allowing the baseline hazard to differ for all state transitions. Probabilities were determined from inverse probability of sampling-weighted multistate survival models for individuals with and without CNV calls. In the survival models, participants were censored if they had not transitioned from a given state by the end of the study or if they had died, immigrated, or otherwise been lost to follow-up.

To study early mortality, we fitted a CPH regression model with age at death (or date of censoring) as the primary outcome and CNV status as an independent variable as coded in the primary analyses. Fertility was studied separately for each sex using a Poisson regression model with the number of offspring as the outcome. Copy number variant status was entered as an independent variable and adjusted for the 5 main iPSYCH disorders. Copy number variant associations with 172 somatic disorders (available for analysis owing to prior indication of an association with mental disorders) was done using the CPH regression model in the entire case cohort and adjusting for the 5 main iPSYCH disorders. Data analysis was conducted from July 1, 2017, to September 7, 2021. All statistical analyses were performed in R, version 3.3.1 (R Foundation for Statistical Computing), and CPH regression models were fitted using the R package survival. With 2-tailed testing, findings were significant at $P < .05$.

Results

This study leveraged the iPSYCH case cohort (eFigure 6 in the Supplement). After initial quality control, the random population cohort included 25 704 individuals and the case sample included 49 250 individuals (eTable 2 in the Supplement).¹⁰ Participants were aged 1 to 32 years (mean, 12.0 [IQR, 6.9] years), 38 662 were male (52.3%), and 35 261 were female (47.7%). The number of samples eligible for analysis following further locus-specific LRR-SD-based quality control differed across the 6 CNV loci (eTable 1 and eMethods in the Supplement). In total, we identified 520 deletions (1 in 142 individuals) and 660 duplications (1 in 122 individuals) in the overall case cohort for the 6 examined loci (eTable 2 in the Supplement).

The prevalence of the identified CNVs in the Danish population was estimated using the random cohort and ranged from 0.021% (95% CI, 0.008%-0.045%) (1:4865) for the 15q13.3 deletion to 0.49% (95% CI, 0.41%-0.59%) (1:203) for the 15q11.2 duplication (Table 1). Duplications were significantly more frequent than deletions at 1q21.1 (ratio, 4.6; $P < 9.1 \times 10^{-4}$) and 15q13.3 (ratio, 3.0; $P < .04$), whereas no significant sex differences in prevalence for any of the CNVs were found (eTable 3 in the Supplement).

For each CNV, we compared observed prevalence rates with those previously reported in the UK Biobank (UKB)¹⁹ and control samples from the largest published meta-analyses of ADHD,²⁰ ASD,²¹ SCZ,²² and BPD²³ studies. A meta-analysis of MDD studies has also been published²⁴ but could not be compared because it did not provide carrier counts by case-control status. This examination yielded 53 pairwise prevalence-rate comparisons: 6 of these found a significantly higher prevalence in the iPSYCH case cohort, involving 4 CNVs (Table 1).

Overall, the carrier rate was higher in the iPSYCH case sample than in unaffected individuals from the iPSYCH cohort for all CNVs except the deletion and duplication at 17p12 (eTable 4 in the Supplement). The case-cohort design and the use of inverse probability of sampling weights enabled us to generate population-unbiased HR estimates of CNV associations with risk for the 5 ascertained psychiatric outcomes (Figure 1 and Figure 2).

Consistent with previous studies, deletions at 1q21.1 and 15q13.3 associated with more diagnoses and showed higher HRs than their respective reciprocal duplications, whereas associations and HR patterns were less clear for other tested loci (Figure 1). We noted that the 17p12 deletion, although not significantly associated with any of the tested disorders, consistently showed an HR less than 1 for each disorder (eTable 4 in the Supplement for HR and 95% CI for each CNV for each diagnosis).

To assess in more detail whether there was a consistent association of decreased vs increased gene dosage, we compared deletion and duplication HRs within each locus (eFigure 2 in the Supplement). This comparison revealed that 15q13.3 and 1q21.1, the loci with the greatest duplication to deletion prevalence ratios, were distinctive for deletions, consistently showing larger HRs for diagnostic outcomes than duplications (eFigure 2 in the Supplement).

Table 1. Prevalence of CNVs and Comparison With Previous Population-Based and Case-Control Studies

LOCUS	Deletions, %				Duplications, %			
	iPsych	Other ^a	OR ^b	P value	iPsych	Other ^a	OR ^b	P value
iPSYCH random sample (n = 19 169-25 431) vs UK Biobank total population (n = 421 268)¹⁹								
1q21.1	0.021	0.027	0.78	.84	0.097	0.042	2.3	6.9 × 10 ^{-4c}
15q11.2	0.44	0.39	1.1	.30	0.49	0.48	1.0	.84
15q13.3	0.021	0.01	2.1	.11	0.062	0.057	1.1	.68
16p11.2	0.052	0.026	2.0	.04	0.11	0.033	3.3	6.5 × 10 ^{-6c}
17p12	0.051	0.056	0.91	.89	0.024	0.029	0.80	.85
17q12	0.023	0.0021	11	4.3 × 10 ^{-4c}	0.023	0.024	0.94	>.99
iPSYCH ADHD cases (n = 12 650-16 427) vs ADHD cases from meta-analysis (n = 8883)²⁰								
1q21.1	0.11	0.079	1.4	.53	0.18	0.15	1.2	.63
15q11.2	0.58	0.42	1.4	.11	0.58	NA	NA	NA
15q13.3	0.070	0.14	0.52	.13	0.13	NA	NA	NA
16p11.2	0.063	0.079	0.80	.79	0.36	0.19	1.9	.03
17p12	≤0.024	0.056	≤0.43	Not significant	≤0.024	NA	NA	NA
17q12	0.048	0	Infinite	.05	0.12	0.068	1.7	.29
iPSYCH ASD cases (n = 11 022-14 145) vs ASD cases from published meta-analysis (n = 2120-4315)²¹								
1q21.1	0.11	0.033	3.2	.33	0.28	0.26	1.1	>.99
15q11.2	0.67	0.094	7.1	3.1 × 10 ^{-4c}	0.46	NA	NA	NA
15q13.3	0.074	0.19	0.39	.11	0.12	NA	NA	NA
16p11.2	0.15	0.42	0.37	.004	0.29	0.39	0.73	.34
17p12	≤0.028	0.094	≤0.30	Not significant	≤0.028	NA	NA	NA
17q12	0.072	0.094	0.76	.67	0.056	NA	NA	NA
iPSYCH SCZ cases (n = 1704-2590) vs SCZ cases from published meta-analysis (n = 21 094)^{22 d}								
1q21.1	≤0.17	0.16	≤1.09	Not significant	≤0.17	0.090	≤1.9	Not significant
15q11.2	0.50	0.45	1.1	.74	0.46	NA	NA	
15q13.3	≤0.16	0.14	≤1.2	Not significant	≤0.16	NA	NA	NA
16p11.2	0	0.052	0	>.99	0.35	0.30	1.2	.64
17p12	≤0.15	0.057	≤2.7	Not significant	≤0.15	NA	NA	NA
17q12	≤0.19	0.019	≤10	Not significant	≤0.19	0.076	≤2.5	Not significant
iPSYCH BPD cases (n = 925-1384) vs BPD cases from published meta-analysis (n = 7250-9129)^{23 d}								
1q21.1	≤0.31	0.033	≤9.4	Not significant	≤0.31	0.099	≤3.2	Not significant
15q11.2	≤0.33	0.17	≤2.0	Not significant	0.42	NA	NA	NA
15q13.3	0	0.025	0	>.99	≤0.30	NA	NA	NA
16p11.2	0	NA	NA		0	0.13	0	.62
17p12	≤0.29	0.049	≤5.9	Not significant	0	NA	NA	NA
17q12	0	0	NA		0	NA	NA	NA
iPSYCH unaffected (n = 18 409-24 408) vs controls from above meta-analyses combined (n = 201 003-358 329)^{20-23e}								
1q21.1	0.022	0.025	0.86	>.99	0.092	0.039	2.3	.001
15q11.2	0.44	0.26	1.7	2.4 × 10 ^{-6c}	0.49	NA	NA	NA
15q13.3	0.021	0.019	1.2	.63	0.060	NA	NA	NA
16p11.2	0.049	0.035	1.4	.31	0.11	0.038	2.9	8.1 × 10 ^{-5c}
17p12	0.053	0.029	1.9	.05	0.025	NA	NA	NA
17q12	0.023	0.0052	4.5	.009	0.023	0.030	0.77	.83

Abbreviations: ADHD, attention-deficit/hyperactivity disorder; ASD, autism spectrum disorder; BPD, bipolar disorder; CNV, copy number variations; iPSYCH, the Lundbeck Foundation Initiative for Integrative Psychiatric Research; MDD, major depressive disorder; NA, not available; OR, odds ratio; SCZ, schizophrenia.

^a None of the comparison studies provided estimates for all the 12 CNVs.

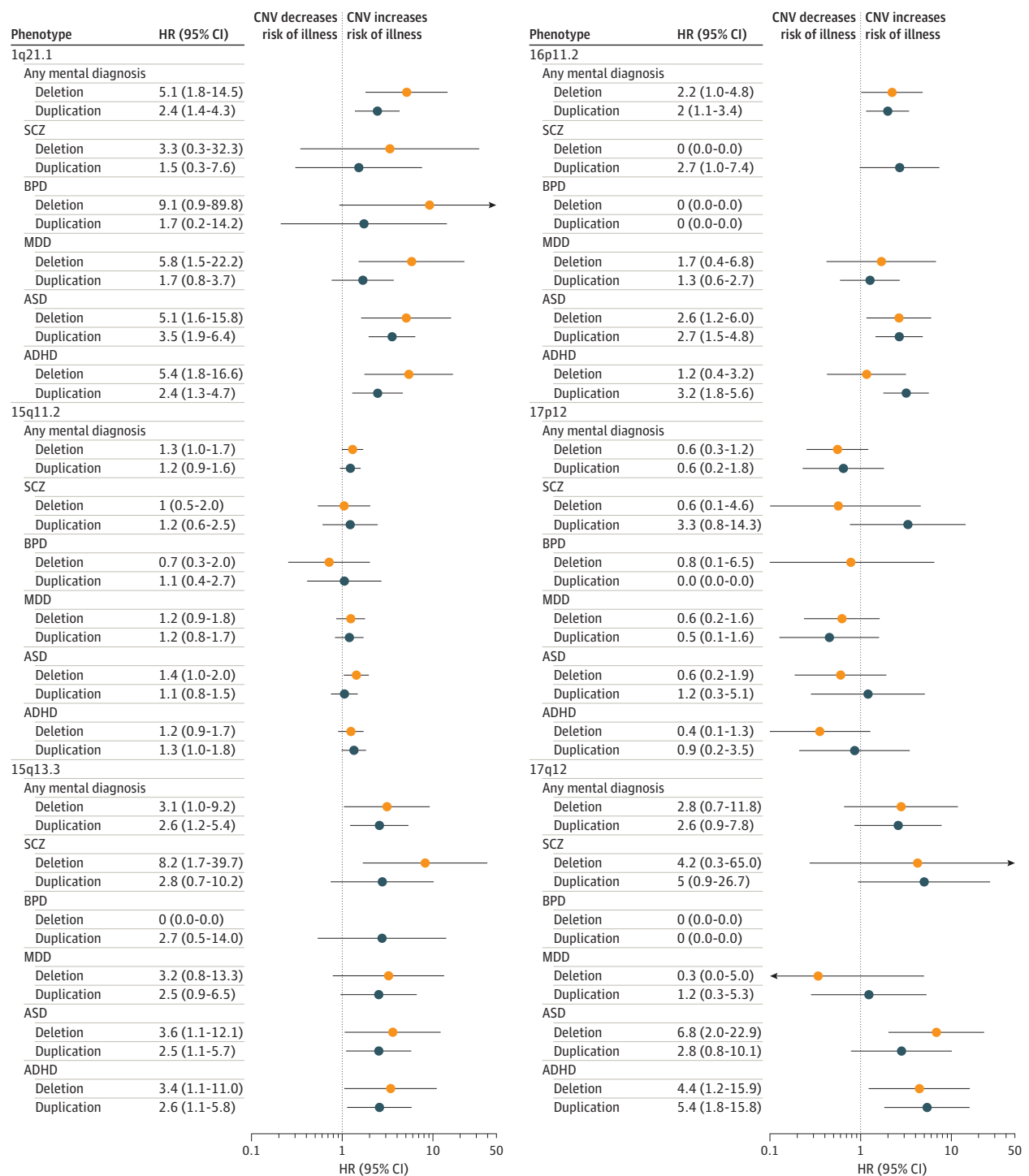
^b An OR greater than 1 indicates higher prevalence in iPSYCH than the comparison study.

^c Comparisons in which difference in prevalence was significant (considering 53 comparisons).

^d For CNVs with fewer than 5 SCZ-affected carriers, prevalence and OR corresponding to less than or equal to 4 carriers are reported.

^e We use prevalence in the unaffected subset of the random sample in comparison with combined control samples of the published meta-analyses for ADHD, ASD, SCZ, and BPD.

Figure 1. Hazard Ratios (HRs) and 95% CIs at Each Studied Copy Number Variation (CNV) Locus for the 5 Ascertained Diagnoses Separately and Combined (Any Mental Diagnosis)

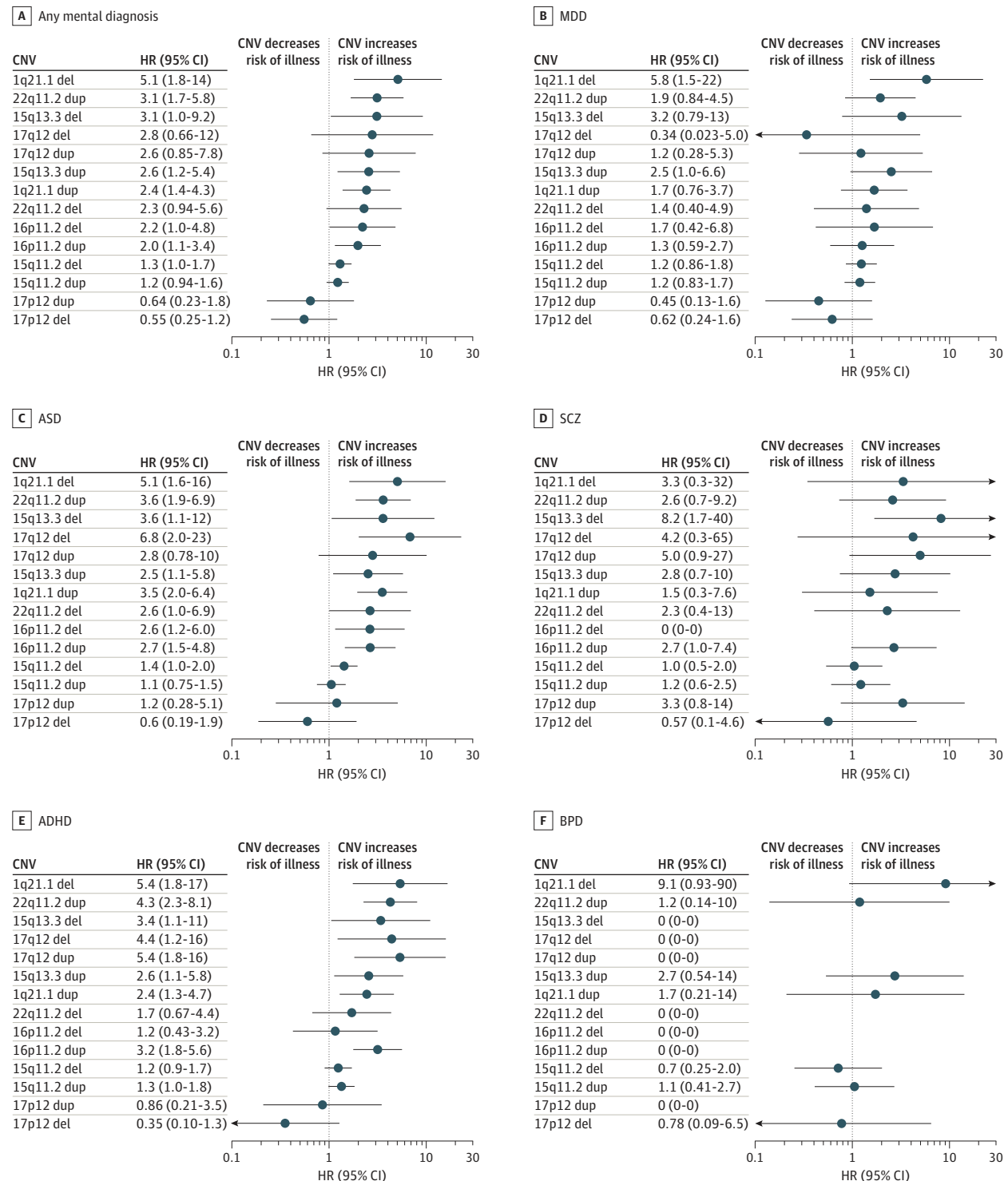


ADHD indicates attention-deficit/hyperactivity disorder; ASD, autism spectrum disorder; BPD, bipolar disorder; MDD, major depressive disorder; and SCZ, schizophrenia. The x-axis is on a logarithmic scale.

We also estimated sex-specific HRs for the 5 ascertained diagnoses. Although distinct baseline hazard functions prevent direct comparisons between sexes, we noted significant HRs

for 1 sex with no indication of an association with the other sex at 5 of the 6 loci for 4 of the 5 disorders. Deletion at 16p11.2 selectively increased the HR in females and duplication at 16p11.2

Figure 2. Hazard Ratios (HRs) and 95% CIs Across the 12 Copy Number Variations (CNVs) for Each of the 5 Ascertained Disorders and Combined (Any Mental Diagnosis)



ADHD indicates attention-deficit/hyperactivity disorder; ASD, autism spectrum disorder; BPD, bipolar disorder; MDD, major depressive disorder; and SCZ, schizophrenia. The x-axis is on a logarithmic scale.

selectively increased the HR in males, most notably for ASD; however, the 1q21.1 deletion was associated with BPD only among males (HR, 17; 95% CI, 1.5-189.3) (eFigure 3 in the Supplement).

To provide a complementary view of neuropsychiatric penetrance for the CNVs, we examined the disease trajectories for individuals with and without CNVs at any of the 6 loci,

calculating the state transition probabilities as a function of age using a multistate survival model.²⁵ The probability of being assigned a diagnosis of ASD, ADHD, or ID before age 32 years was increased among carriers for all of the tested CNVs, except the 17p12 deletion (1.1% vs 2.9% of individuals without CNV) with the highest probabilities seen for the 1q21.1 deletion (10.7%) and 17q12 duplication (10.6%) (Figure 3). A considerable increase in the risk of ID comorbid with ASD or ADHD was observed for deletions and duplications at 15q13.3 and 17q12; a minor increase was seen for deletions and duplications at 15q11.2 and 16p11.2 and for the duplication at 17p12. Comorbidity between ASD and ADHD was substantially increased for the 1q21.1 deletion and 16p11.2 duplication and slightly increased for duplications at 1q21.1 and 17q12 and deletion of 15q13.3 (Figure 3). We also assessed weighted hazards for epilepsy and ID and found HRs for epilepsy to be significantly increased for 16p11.2 (deletion and duplication) and in carriers of the 17p12 duplication (eTable 4 in the Supplement).

The observations detailed herein hint at the possibility of differential psychiatric risk profiles across CNVs. To test this hypothesis, we ranked the CNVs according to their overall pathogenicity (ie, by HR for any mental diagnosis) (Figure 1), including the 22q11.2 deletion and duplication in these analyses that have also been previously examined (Figure 2).⁹ Although the rank order of the 14 CNVs across all diagnoses was largely recapitulated for ASD and ADHD, differences were observed for MDD, SCZ, and BPD. In fact, increased risk of MDD, SCZ, and BPD was mainly observed for CNVs that also conferred a high risk of ASD and ADHD. Again, 17p12 deletion carriers consistently showed a lower risk of disease than those without any of the tested CNVs.

Subsequent analyses found reduced fertility (ie, reduced number of offspring until age 32 years) only in male carriers of 15q11.2 duplications, and a trend for reduced fertility among male carriers of 1q21.1 duplication and female carriers of the 16p11.2 deletion (eFigure 4, eTable 5 in the Supplement). Early mortality hazards (age, 1-32 years) could be estimated only for CNVs at 15q11.2 and were not significantly increased (duplication [$n < 5$]: mean [SD] age, 18.9 [2.2] years; HR, 1.23; 95% CI, 0.39-3.87; $P = .72$ and deletion [$n < 5$]: mean [SD] age, 23.0 [5.0] years; HR, 1.44; 95% CI, 0.47-4.45; $P = .52$).

In addition, to explore CNV association with specific somatic disorders, we selected disorders that were observed in at least 2 carriers of 1 of the 12 CNVs ($n = 18$) (eTable 6 in the Supplement) for formal comparison in the case cohort. We observed 9 nominally significant associations involving 7 CNVs (Table 2; eFigure 5 in the Supplement). Most markedly, the 15q13.3 deletion was associated with thyroiditis (HR, 44.8; 95% CI, 10.8-185.4; $P < .001$), 16p11.2 duplication with febrile seizures (HR, 2.7; 95% CI, 1.5-4.9; $P < .001$), and 17p12 duplication with gastrointestinal infection (HR, 4.4; 95% CI, 1.9-9.9; $P < .001$). Also, gastrointestinal infections and febrile seizures were significantly overrepresented among carriers of the 17q12 deletion (eFigure 5 in the Supplement).

Discussion

This is, to our knowledge, the first unbiased, population-based, comparative study of the prevalence of psychiatric illness in individuals with CNVs at the 6 tested loci. For these CNVs, we estimated age-dependent and sex-specific HRs for 5 major psychiatric disorders in a sample truly representative of the Danish population. To our knowledge, we provide the first evidence that the 1q21.1 deletion confers risk for MDD and, in men only, for BPD. Although our study was well powered for MDD ($n = 20\,084$), the observation of risk of BPD was unexpected given the considerably smaller sample size ($n = 1272$) and possible only owing to the large effect size.

We did not observe an increased risk of mental disorders in individuals carrying the 17p12 deletion; rather, for all ascertained diagnoses, the 17p12 deletion showed HRs below 1 and was ranked as less pathogenic than all other CNVs. Our findings support previous reports of the 16p11.2 duplication increasing risk for ASD,¹⁵ ADHD,²⁰ and epilepsy^{26,27} (HR, 2.7-6.6), and of a high risk of SCZ (HR, 8.2) conferred by the 15q13.3 deletion.⁶ Our risk estimates for SCZ, ASD, and ADHD were generally lower across the examined CNVs than previously reported.²⁰⁻²² The main exceptions from this trend are the 1q21.1 deletion (ASD and ADHD) and 17q12 duplication (ADHD), conferring a 2- to 3-fold higher risk in our study.

Although these discoveries may in part be associated with large, homogeneous patient samples, in particular for ASD, ADHD, and MDD, the overall trend for reduced risk estimates relative to previous studies seems in large part to be associated with a higher CNV prevalence in unaffected individuals from the iPSYCH case cohort compared with control groups in the previous studies²⁰⁻²³ (the same trend was seen in the comparison of prevalence between the iPSYCH cohort and the UKB database¹⁹) (Table 1). This difference in prevalence, and subsequently in risk estimates, most likely arises from the iPSYCH case cohort being ascertained in a different way than the samples of the other studies. In fact, our analysis provides estimates as if we had analyzed the whole Danish population under the assumption that quality control is independent of CNV carrier status and diagnosis. In contrast, the UKB sample is based on voluntary participation, with participants aged 40 to 69 years at the time of recruitment with a response rate of only 5.5%.²⁸ In addition to the survival bias caused by the participants' relatively older age at recruitment, the UKB sample is underrepresentative of individuals who smoke, have a high rate of alcohol use, and are of low socioeconomic status²⁹—all features associated with mental illness. Clinical case-control studies also rely on voluntary participation, with control participants often having been screened for a history of mental disorders.

The population-representative case-cohort design of the present study allowed us to compare and rank CNVs across multiple diagnoses as well as between sexes in a single and consistent set of analyses. First, within-locus analyses revealed more duplications than deletions (ratio 1.33:1); this ratio increased further (to 1.88:1) when discounting the relatively common 15q11.2 CNVs. These ratios are very similar to those seen

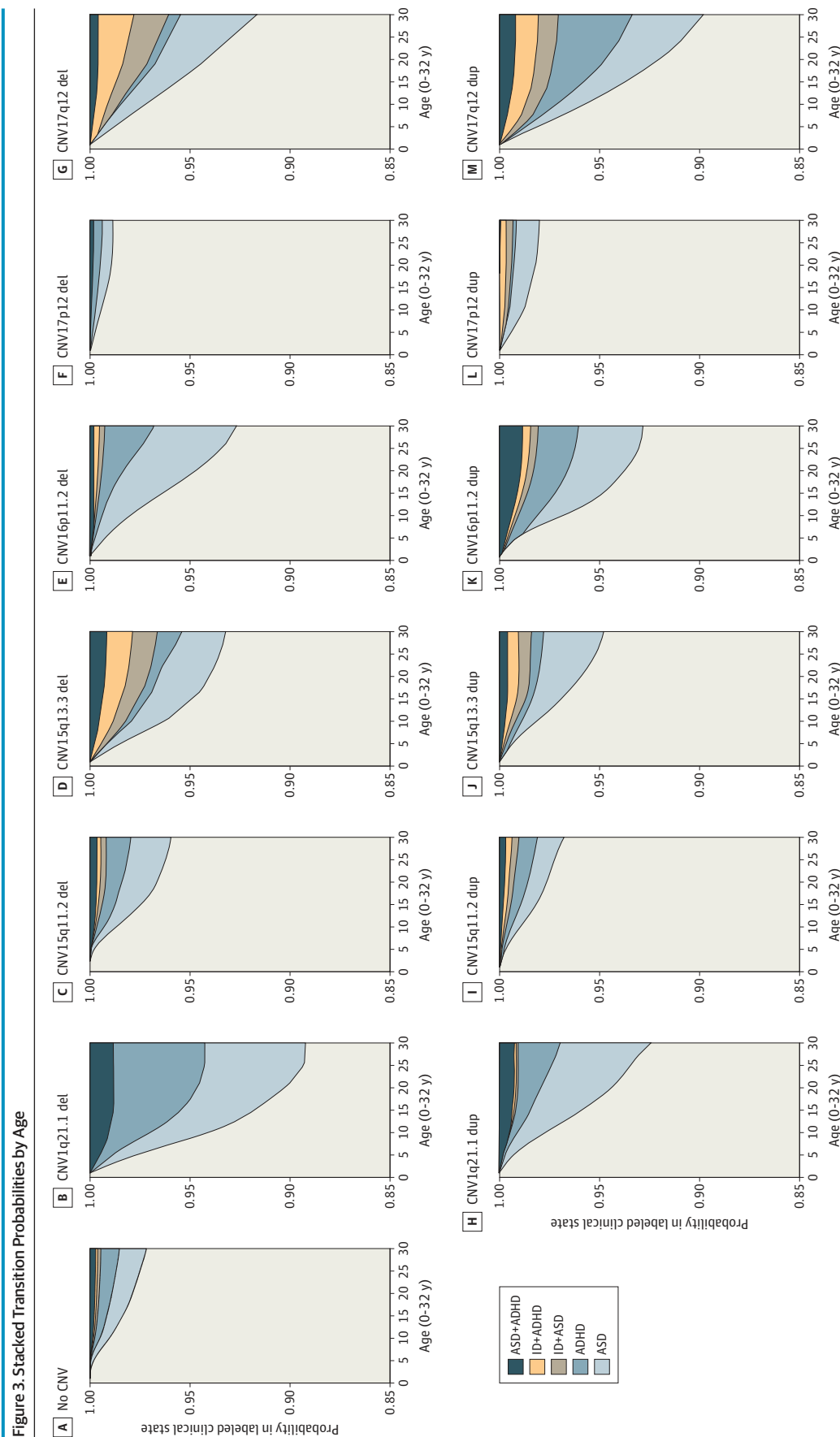


Figure 3. Stacked Transition Probabilities by Age

Probabilities by age (0-32 years) were determined from inverse probability of sampling-weighted multistate survival models for individuals with and without copy number variation (CNV). In the survival models, participants were censored if they had not transitioned from a given state by the end of the study or if they had died, emigrated, or were otherwise lost to follow-up. The distance between 2 adjacent lines represents the probability of being in a given state (eg, attention-deficit/hyperactivity disorder [ADHD] or autism spectrum disorder [ASD]) in the unexposed sample (no deletion or duplication), in carriers of a deletion, and in carriers of the corresponding duplication. ID indicates intellectual disability.

Table 2. Somatic Disorders Observed Significantly More Frequently in CNV Carriers Than the General Population

Locus	Type	Diagnosis	HR (95% CI)	P value	Without diagnosis, No.		With diagnosis, No.	
					Noncarrier	Carrier	Noncarrier	Carrier
1q21.1	Deletion	Infection (otitis)	2.41 (1.19-4.9)	.02	63 588	38	6493	9
15q11.2	Deletion	Juvenile arthritis	4.26 (1.76-10.38)	.001	65 426	347	225	5
		Thrombocytopenic purpura	4.73 (1.16-19.29)	.03	65 572	350	79	<5
15q13.3	Deletion	Thyroiditis	44.78 (10.81-185.4)	2.12E-07	71 756	35	83	<5
16p11.2	Deletion	Asthma	2.378 (1.19-4.74)	.014	49 461	30	6340	9
	Deletion	Gastrointestinal (infection)	2.14 (0.10-4.6)	.05	49 598	31	6203	8
	Duplication	Febrile seizures	2.71 (1.5-4.88)	<.001	53 319	97	2482	12
	Duplication	Otitis (infection)	1.73 (1.04-2.85)	.03	50 542	91	5259	18
17p12	Duplication	Infection (gastrointestinal)	4.4 (1.96-9.88)	<.001	66 821	11	8451	6
	Duplication	Syncope	4.23 (1.32-13.59)	.02	72027	14	3245	<5
17q12	Deletion	Infection (gastrointestinal)	3.41 (1.46-7.9)	.005	58 048	14	7312	7
		Febrile seizures	4.21 (1.42-12.5)	.009	62 478	17	2882	<5
	Duplication	Febrile seizures	3.07 (1.16-8.09)	.02	62 478	31	2882	5

Abbreviations: CNV, copy number variation; HR, hazard ratio.

in the UKB (1.46:1 for the 6 loci, and 1.67:1 without 15q11.2) and consistent with previous observations that deletions are more pathogenic and therefore less viable and/or less frequently transmitted than duplications. However, among the loci investigated, we saw no clear trend for more reduced fertility among deletion carriers than duplication carriers (eTable 5 in the Supplement), nor did we find that deletions consistently appeared to confer a higher risk of mental disorders than duplications (eFigure 2 in the Supplement). Also, we did not observe increased mortality among deletion carriers, although the mortality as well as fertility analyses were limited in power by the younger age of the case cohort (1-32 years). It is notable though that duplication-to-deletion ratios were greatest for the 15q13.3 and 1q21.1 loci because these are also the 2 loci in which disparity in psychiatric risk between deletions and duplications is greatest and consistently associated with greater risk in the deletions across diagnoses (eFigure 2 in the Supplement).

Second, the consistent rank order of HRs across diagnoses suggests that the examined CNVs influence the risk of ASD and ADHD similarly and supports the notion that CNVs perturb neurodevelopmental processes common to different mental disorders. However, notwithstanding wide 95% CIs in HR estimates, our findings suggest that some CNVs may pose risks for particular diagnostic outcomes that are out of keeping with the risk they impose for others. Most notably, 17q12 deletion and duplication appear similarly penetrant across most index diagnoses, except MDD, where we saw no trend for increased risk. The hazard rank order also suggests that CNV-conferred risk of adult-onset disorders is linked with a high risk of childhood-onset disorders, possibly informing on part of the clinical heterogeneity of adult mental disorders.

Third, we provide preliminary evidence that CNV-conferred risk of psychopathology differs between sexes, with examples involving most of the tested CNVs and disorders. The finding that CNVs at the 16p11.2 locus affect risk of ASD differently, with increased hazards conferred in women by the deletion and in men by the duplication, is particularly inter-

esting given the considerable difference in the baseline hazard for ASD between sexes and may represent an experimental framework important for the identification of disease mechanisms in ASD and more generally in mental disorders.

To gain clinically relevant insight into the developmental dynamism of mental disorders between CNVs, we considered transition probabilities between diagnoses by age, including ID, although this diagnosis was not directly ascertained in the case cohort. The risk of comorbid ID seemed locus- rather than CNV-dependent, being prominent at the 15q13.3 and 17q12 loci and negligible at the otherwise pathogenic 1q21.1 locus. The age dependence of diagnoses and comorbidities in CNV carriers also hints that ASD and ADHD are diagnosed earlier if comorbid with ID. Although this age dependence might reflect more severe forms of the 2 disorders, it could also result from observation bias following the diagnosis of ID, which would suggest that ASD and ADHD in the absence of ID could be diagnosed earlier in life, benefiting affected children and their parents.

We also tested for enrichment of specific somatic disorders in CNV carriers, correcting for mental disorders. The most significant enrichment was found for thyroiditis among 15q13.3 deletion carriers. To date, case series of 15q13.3 deletion syndrome have documented 1 case of congenital hypothyroidism,³⁰ but we are otherwise not aware of a previously documented association between 15q13.3 and thyroid disease. It remains to be seen whether this association is replicable in other cohorts and whether it reflects a direct primary outcome of haploinsufficiency at 15q13.3 or is a secondary consequence of other factors associated with this CNV (eg, medication treatment). We also found significant associations with different CNVs of other immune-related disorders, including infections, juvenile arthritis, and thrombocytopenic purpura. Although some of these outcomes have been reported to be associated broadly with mental disorders³¹ and may reflect secondary effects of illness, our finding may point toward a primary CNV association with immune system function.

Although the iPSYCH case cohort is still relatively young and therefore not yet well powered to estimate CNV-specific associations with fertility, we saw a trend for reduced fertility among female carriers of the 16p11.2 deletion and male carriers of the 1q21.1 and 15q11.2 duplication. We noted that the point estimates of the former 2 CNVs suggest a greater association with fertility than that of the more common 15q11.2 duplication. A previous study on fertility of CNV carriers found the strongest reduction for the 16p11.2 deletion, although more pronounced in male than female carriers, consistent with this CNV possibly being under particularly strong negative selection (the previous study did not test the 15q11.2 duplication).³²

Limitations

This study has limitations. Although the uniqueness of the iPSYCH case cohort lies in the genuine representativeness of the Danish population, allowing for estimation of age-dependent absolute risk, our findings must be considered in light of several limitations. First, we were not able to capture subdiagnostic outcomes, which could potentially lead to an underestimation of CNV penetrance for clinically significant psychopathologic outcomes. Second, although our study is the largest population-based analysis of CNV penetrance, the base rate of CNVs and disease outcomes is such that 95% CIs around our point estimates are still wide. Third, although we applied rigorous quality control and validation of CNV calls, it is possible that errors in genotyping and CNV calling could be limiting our precision in risk

estimation. Future expansion of population-based studies will be an important step forward in this regard.

Conclusions

Notwithstanding the abovementioned limitations, our study updates the understanding of genetic risk for psychopathologic disorders in several directions. First, we showed that population-based analyses generally increase prevalence estimates for most studied CNVs and decrease penetrance estimates for most studied outcomes. Second, we noted a trend toward 17p12 deletion, potentially imparting some protection against risk for psychopathologic disorders. This finding warrants careful targeted follow-up because, if confirmed, it may ultimately point toward new therapeutic opportunities. Third, comparative analysis indicated that the balance of risk between deletions and duplications can vary between loci. Fourth, we observed a largely consistent rank ordering of CNVs by risk across various diagnostic outcomes, with notable exceptions. Fifth, we found evidence that the penetrance of a CNV for psychopathologic disorders may also be modulated by carrier sex, offering potential inroads into the broader phenomenon of sex differences in the prevalence of psychiatric disorders. Sixth, by looking beyond psychiatric outcomes, we noted that some of the CNVs studied appeared to modify risk for selected somatic disorders, which identified new disease associations for future research.

ARTICLE INFORMATION

Accepted for Publication: October 4, 2021.

Published Online: November 24, 2021.
doi:10.1001/jamapsychiatry.2021.3392

Correction: This article was corrected on January 5, 2022, to fix a numeric error in Figure 1.

Author Affiliations: Institute of Biological Psychiatry, Mental Health Services, Copenhagen University Hospital, Copenhagen, Denmark (Calle Sánchez, Helenius, Montalbano, Ingason, Werge); The Lundbeck Foundation Initiative for Integrative Psychiatric Research, Copenhagen and Aarhus, Denmark (Calle Sánchez, Helenius, Børnglum, Bybjerg-Grauholm, Pedersen, Hougaard, Børnglum, Nordentoft, Mors, Mortensen, Montalbano, Thompson, Ingason, Werge); Section of Neonatal Screening and Hormones, Department of Clinical Biochemistry, Immunology and Genetics, Statens Serum Institut, Copenhagen, Denmark (Bybjerg-Grauholm, Hougaard); National Centre for Register-based Research, Aarhus University, Aarhus, Denmark (Pedersen, Mortensen); Centre for Integrated Register-based Research at Aarhus University, Aarhus, Denmark (Pedersen); Big Data Centre for Environment and Health, Aarhus University, Aarhus, Denmark (Pedersen); Department of Biomedicine and iSEQ-Centre for Integrative Sequencing, Aarhus University, Denmark (Børnglum, Mortensen); Centre for Genomics and Personalized Medicine, Aarhus University, Denmark (Børnglum); Copenhagen Research Centre for Mental Health, Mental Health Centre Copenhagen, Copenhagen University Hospital, Copenhagen, Denmark (Nordentoft); Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark (Nordentoft,

Werge); Psychosis Research Unit, Aarhus University Hospital, Risskov, Denmark (Mors); Department of Neurology, University of California, Los Angeles (Geschwind); Department of Human Genetics, University of California, Los Angeles (Geschwind); Center for Autism Research and Treatment, Semel Institute, David Geffen School of Medicine at UCLA, University of California, Los Angeles (Geschwind); Center for Human Development, University of California, San Diego (Geschwind); Program in Neurobehavioral Genetics, Semel Institute, David Geffen School of Medicine at UCLA, University of California Los Angeles (Geschwind); Section on Developmental Neurogenetics, Human Genetics Branch, National Institute of Mental Health Intramural Research Program, Bethesda, Maryland (Raznahan); Herbert Wertheim School of Public Health and Human Longevity, University of California, San Diego, La Jolla, California (Thompson); Lundbeck Foundation Center for GeoGenetics, GLOBE Institute, University of Copenhagen, Copenhagen, Denmark (Ingason, Werge).

Author Contributions: Drs Ingason and Werge contributed equally, had full access to all of the data in the study, and take responsibility for the integrity of the data and the accuracy of the data analysis.
Concept and design: Calle Sánchez, Helenius, Pedersen, Hougaard, Nordentoft, Geschwind, Raznahan, Ingason, Werge.
Acquisition, analysis, or interpretation of data: Calle Sánchez, Helenius, Bybjerg-Grauholm, Pedersen, Børnglum, Mors, Mortensen, Montalbano, Thompson, Ingason, Werge.
Drafting of the manuscript: Calle Sánchez, Raznahan, Thompson, Ingason, Werge.

Critical revision of the manuscript for important intellectual content: Calle Sánchez, Helenius, Bybjerg-Grauholm, Pedersen, Hougaard, Børnglum, Nordentoft, Mors, Mortensen, Geschwind, Montalbano, Ingason, Werge.
Statistical analysis: Calle Sánchez, Helenius, Montalbano, Raznahan, Thompson, Werge.
Obtained funding: Calle Sánchez, Hougaard, Børnglum, Nordentoft, Mors, Mortensen, Ingason, Werge.
Administrative, technical, or material support: Calle Sánchez, Bybjerg-Grauholm, Pedersen, Hougaard, Mors, Mortensen, Werge.
Supervision: Helenius, Børnglum, Geschwind, Ingason, Werge.

Conflict of Interest Disclosures: Dr Børnglum reported grants from the Lundbeck Foundation during the conduct of the study. Dr Ingason reported grants from Lundbeck Foundation, grants from National Institutes of Health, grants from Aarhus University, and grants from Copenhagen University Hospital during the conduct of the study. Dr Werge reported grants from Lundbeck Foundation and grants from National Institutes of Health during the conduct of the study. No other disclosures were reported.

Funding/Support: The Initiative for Integrative Psychiatric Research (iPSYCH) was supported by grants from the Lundbeck Foundation (R165-2013-15320, R155-2014-1724, and R248-2017-2003), the National Institutes of Health (R01MH124789-01), and the universities and university hospitals of Aarhus and Copenhagen, Denmark. Mr Calle Sánchez acknowledges support from the Lundbeck Foundation (Lundbeck Foundation R286-2018-1510). This research was

conducted using the Danish National Biobank resource, supported by the Novo Nordisk Foundation.

Role of the Funder/Sponsor: The funding organizations had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Additional Information: The iPSYCH was approved by the Danish Scientific-Ethical Committee.

REFERENCES

- Kirov G. CNVs in neuropsychiatric disorders. *Hum Mol Genet.* 2015;24(R1):R45-R49. doi:10.1093/hmg/ddv253
- Coelewij L, Curtis D. Mini-review: update on the genetics of schizophrenia. *Ann Hum Genet.* 2018;82(5):239-243. doi:10.1111/ahg.12259
- Lowther C, Costain G, Baribeau DA, Bassett AS. Genomic disorders in psychiatry—what does the clinician need to know? *Curr Psychiatry Rep.* 2017;19(11):82. doi:10.1007/s11920-017-0831-5
- McDonald-McGinn DM, Sullivan KE, Marino B, et al. 22q11.2 deletion syndrome. *Nat Rev Dis Primers.* 2015;1:15071. doi:10.1038/nrdp.2015.71
- Murphy KC, Jones LA, Owen MJ. High rates of schizophrenia in adults with velo-cardio-facial syndrome. *Arch Gen Psychiatry.* 1999;56(10):940-945. doi:10.1001/archpsyc.56.10.940
- Stefansson H, Rujescu D, Cichon S, et al; GROUP. Large recurrent microdeletions associated with schizophrenia. *Nature.* 2008;455(7210):232-236. doi:10.1038/nature07229
- International Schizophrenia Consortium. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature.* 2008;455(7210):237-241. doi:10.1038/nature07239
- Vassos E, Collier DA, Holden S, et al. Penetrance for copy number variants associated with schizophrenia. *Hum Mol Genet.* 2010;19(17):3477-3481. doi:10.1093/hmg/ddq259
- Olsen L, Sparsø T, Weinsheimer SM, et al. Prevalence of rearrangements in the 22q11.2 region and population-based risk of neuropsychiatric and developmental disorders in a Danish population: a case-cohort study. *Lancet Psychiatry.* 2018;5(7):573-580. doi:10.1016/S2215-0366(18)30168-8
- Pedersen CB, Bybjerg-Grauholm J, Pedersen MG, et al. The iPSYCH2012 case-cohort sample: new directions for unravelling genetic and environmental architectures of severe mental disorders. *Mol Psychiatry.* 2018;23(1):6-14. doi:10.1038/mp.2017.196
- Kirov G, Grozeva D, Norton N, et al; International Schizophrenia Consortium; Wellcome Trust Case Control Consortium. Support for the involvement of large copy number variants in the pathogenesis of schizophrenia. *Hum Mol Genet.* 2009;18(8):1497-1503. doi:10.1093/hmg/ddp043
- Moreno-De-Luca D, Mulle JG, Kaminsky EB, et al; SGENE Consortium; Simons Simplex Collection Genetics Consortium; GeneSTAR. Deletion 17q12 is a recurrent copy number variant that confers high risk of autism and schizophrenia. *Am J Hum Genet.* 2010;87(5):618-630. doi:10.1016/j.ajhg.2010.10.004
- Szatkiewicz JP, O'Dushlaine C, Chen G, et al. Copy number variation in schizophrenia in Sweden. *Mol Psychiatry.* 2014;19(7):762-773. doi:10.1038/mp.2014.40
- Bernier R, Steinman KJ, Reilly B, et al; Simons VIP consortium. Clinical phenotype of the recurrent 1q21.1 copy-number variant. *Genet Med.* 2016;18(4):341-349. doi:10.1038/gim.2015.78
- Weiss LA, Shen Y, Korn JM, et al; Autism Consortium. Association between microdeletion and microduplication at 16p11.2 and autism. *N Engl J Med.* 2008;358(7):667-675. doi:10.1056/NEJMoa075974
- McCarthy SE, Makarov V, Kirov G, et al; Wellcome Trust Case Control Consortium. Microduplications of 16p11.2 are associated with schizophrenia. *Nat Genet.* 2009;41(11):1223-1227. doi:10.1038/ng.474
- Schmidt M, Schmidt SAJ, Sandegaard JL, Ehrenstein V, Pedersen L, Sørensen HT. The Danish National Patient Registry: a review of content, data quality, and research potential. *Clin Epidemiol.* 2015;7:449-490. doi:10.2147/CLEP.S91125
- Barlow WE, Ichikawa L, Rosner D, Izumi S. Analysis of case-cohort designs. *J Clin Epidemiol.* 1999;52(12):1165-1172. doi:10.1016/S0895-4356(99)00102-X
- Crawford K, Bracher-Smith M, Owen D, et al. Medical consequences of pathogenic CNVs in adults: analysis of the UK Biobank. *J Med Genet.* 2019;56(3):131-138. doi:10.1136/jmedgenet-2018-105477
- Gudmundsson OO, Walters GB, Ingason A, et al. Attention-deficit hyperactivity disorder shares copy number variant risk with schizophrenia and autism spectrum disorder. *Transl Psychiatry.* 2019;9(1):258. doi:10.1038/s41398-019-0599-y
- Malhotra D, Sebat J. CNVs: harbingers of a rare variant revolution in psychiatric genetics. *Cell.* 2012;148(6):1223-1241. doi:10.1016/j.cell.2012.02.039
- Marshall CR, Howrigan DP, Merico D, et al; Psychosis Endophenotypes International Consortium; CNV and Schizophrenia Working Groups of the Psychiatric Genomics Consortium. Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects. *Nat Genet.* 2017;49(1):27-35. doi:10.1038/ng.3725
- Green EK, Rees E, Walters JTR, et al. Copy number variation in bipolar disorder. *Mol Psychiatry.* 2016;21(1):89-93. doi:10.1038/mp.2014.174
- Kendall KM, Rees E, Bracher-Smith M, et al. Association of rare copy number variants with risk of depression. *JAMA Psychiatry.* 2019;76(8):818-825. doi:10.1001/jamapsychiatry.2019.0566
- Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. *Stat Med.* 2007;26(11):2389-2430. doi:10.1002/sim.2712
- Reinthal EM, Lal D, Lebon S, et al; 16p11.2 European Consortium; EPICURE Consortium; EuroEPINOMICS Consortium. 16p11.2 600 kb Duplications confer risk for typical and atypical rolandic epilepsy. *Hum Mol Genet.* 2014;23(22):6069-6080. doi:10.1093/hmg/ddu306
- Niarchou M, Chawner SJRA, Doherty JL, et al. Psychiatric disorders in children with 16p11.2 deletion and duplication. *Transl Psychiatry.* 2019;9(1):8. doi:10.1038/s41398-018-0339-8
- Swanson JM. The UK Biobank and selection bias. *Lancet.* 2012;380(9837):110. doi:10.1016/S0140-6736(12)61179-9
- Fry A, Littlejohns TJ, Sudlow C, et al. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. *Am J Epidemiol.* 2017;186(9):1026-1034. doi:10.1093/aje/kwx246
- Lowther C, Costain G, Stavropoulos DJ, et al. Delineating the 15q13.3 microdeletion phenotype: a case series and comprehensive review of the literature. *Genet Med.* 2015;17(2):149-157. doi:10.1038/gim.2014.83
- Nudel R, Wang Y, Appadurai V, et al. A large-scale genomic investigation of susceptibility to infection and its association with mental disorders in the Danish population. *Transl Psychiatry.* 2019;9(1):283. doi:10.1038/s41398-019-0622-3
- Stefansson H, Meyer-Lindenberg A, Steinberg S, et al. CNVs conferring risk of autism or schizophrenia affect cognition in controls. *Nature.* 2014;505(7483):361-366. doi:10.1038/nature12818