# Similar Rates of Deleterious Copy Number Variants in Early-Onset Psychosis and Autism Spectrum Disorder

Catherine A. Brownstein, Ph.D., Elise Douard, M.Sc., Josephine Mollon, Ph.D., Richard Smith, Ph.D., Margaret A. Hojlo, B.A., Ananth Das, B.A., Maria Goldman, B.A., Emily Garvey, B.A., Kristin Cabral, B.S., Jianqiao Li, B.S., Joshua Bowen, B.S., Abhijit S. Rao, B.S.A., Casie Genetti, M.S., C.G.C., Devon Carroll, A.P.R.N.-B.C., Emma E. M. Knowles, Ph.D., Emma Deaso, B.A., Pankaj B. Agrawal, M.D., Alan H. Beggs, Ph.D., Eugene D'Angelo, Ph.D., Laura Almasy, Ph.D., Aaron Alexander-Bloch, M.D., Ph.D., Zohra Saci, Ph.D., Clara A. Moreau, Ph.D., Guillaume Huguet, Ph.D., Anthony J. Deo, M.D., Ph.D., Sébastien Jacquemont, M.D., David C. Glahn, Ph.D., Joseph Gonzalez-Heydrich, M.D.

**Objective:** Copy number variants (CNVs) are strongly associated with neurodevelopmental and psychotic disorders. Early-onset psychosis (EOP), where symptoms appear before 18 years of age, is thought to be more strongly influenced by genetic factors than adult-onset psychotic disorders. However, the prevalence and effect of CNVs in EOP is unclear.

**Methods:** The authors documented the prevalence of recurrent CNVs and the functional impact of deletions and duplications genome-wide in 137 children and adolescents with EOP compared with 5,540 individuals with autism spectrum disorder (ASD) and 16,504 population control subjects. Specifically, the frequency of 47 recurrent CNVs previously associated with neurodevelopmental and neuropsychiatric illnesses in each cohort were compared. Next, CNV risk scores (CRSs), indices reflecting the dosage sensitivity for any gene across the genome that is encapsulated in a deletion or duplication separately, were compared between groups.

**Results:** The prevalence of recurrent CNVs was significantly higher in the EOP group than in the ASD (odds ratio=2.30) and control (odds ratio=5.06) groups. However, the difference between the EOP and ASD groups was attenuated when EOP participants with co-occurring ASD were excluded. CRS was significantly higher in the EOP group compared with the control group for both deletions (odds ratio=1.30) and duplications (odds ratio=1.09). In contrast, the EOP and ASD groups did not differ significantly in terms of CRS.

**Conclusions:** Given the high frequency of recurrent CNVs in the EOP group and comparable CRSs in the EOP and ASD groups, the findings suggest that all children and adolescents with a psychotic diagnosis should undergo genetic screening, as is recommended in ASD.

*Am J Psychiatry 2022; 179:853–861; doi: 10.1176/appi.ajp.21111175*

Rare copy number variants (CNVs) are deletions and duplications of genomic segments, some with high relative risk for psychotic disorders like schizophrenia (1–3). Recurrent CNVs are relatively common (1 in 10,000 or more) and generally occur due to nonallelic homologous recombination, resulting in similar or identical mutations in unrelated individuals (4, 5), and they are found in approximately 2% of those with adult-onset idiopathic schizophrenia (6, 7). Individuals with childhood-onset schizophrenia whose symptoms begin before age 13 have a significantly higher recurrent CNV burden (10% vs. 2%–6%; $p<0.0001$) relative to those with adult-onset illness, suggesting a greater genetic component in the childhood form of the disorder (8, 9). However, childhood-onset schizophrenia is rare (10), and few cohorts have been genetically characterized to date (8, 9). Moreover, only half of children and adolescents with a psychiatric diagnosis that includes prominent psychotic features meet the strict criteria for schizophrenia (11), and childhood diagnoses often change over the course of development (12). Consequently, there is considerable interest in understanding the genetic underpinnings of the more inclusive early-onset psychosis (EOP) categorization, which captures psychotic symptomatology in various diagnoses. EOP, defined as any psychiatric diagnosis with pronounced psychotic symptoms with onset before age 18, is associated with lower premorbid psychosocial function, more hospitalizations, poorer cognitive functioning, and worse overall prognosis than adult-onset illness (11, 13). Yet, functional outcomes are highly variable in youths with EOP (11, 12), and CNV status has been shown to influence these outcomes (14). Although genomic information could help to disentangle

*See related features:* **Editorial** *by Dr. Skuse (p. 798) and* **Video** *by Dr. Pine (online)*

the clinical heterogeneity in EOP, the genetic architecture of EOP is largely unknown.

Establishing the burden of recurrent CNVs in EOP is an important first step in characterizing the genetic architecture of this extreme phenotype (3). Documenting the frequencies of recurrent CNVs, mutations that occur at high enough rates in the population to foster their individual study, would facilitate the comparison of individuals with EOP with unaffected individuals and individuals with other neurodevelopmental disorders. However, ~90% of CNVs identified in the clinic are nonrecurrent and are too rare (i.e., insufficient copies) for individual association studies to be practical (15). Recently, we developed a strategy to estimate an individual's genome-wide CNV burden by deriving a single aggregate CNV risk score (CRS) that reflects the probability of intolerance to haploinsufficiency of each gene encapsulated by every CNV across the genome, regardless of the mutation's population prevalence (16–18). The CRS is a scalar value that is broadly analogous to the polygenic risk score (PRS), except that while the PRS reflects an individual's liability for an illness based on common genetic variation, the CRS reflects the dosage sensitivity for all genes across the genome that is encapsulated in a deletion or duplication separately. Here, the CRS was estimated with the loss-of-function observed/expected upper bound fraction (LOEUF) score (19). The LOEUF is calculated by comparing the observed and expected number of loss-of-function mutations for a given gene in a reference population (19). Low LOEUF scores indicate strong selection against predicted loss-of-function variation in a gene, and high LOEUF scores indicate relatively higher tolerance to inactivation. Thus, LOEUF scores provide a method for documenting the biological ramifications of individual genes and inferring pathobiology (16, 18, 20). Our CRS measure has been successfully used to model autism spectrum disorder (ASD) (18) and general intelligence (16, 17). Applying this method to an EOP sample enables us to model genome-wide dosage sensitivity in EOP and directly compare our index with ASD and unselected control cohorts, even in a context where the individual CNVs are extremely rare.

Findings that 5%–15% of children with ASD carry a deleterious genetic mutation (21) have led organizations like the American Academy of Pediatrics to recommend that individuals presenting with ASD symptoms undergo genomic screening (22, 23). These established guidelines involve routine use of chromosomal microarrays to document CNVs (24). The burden of recurrent CNVs was similar for ASD and a small childhood-onset schizophrenia cohort (8), suggesting that these disorders have comparable genetic architectures and should be subject to similar genetic screening approaches. Correspondingly, if the burden of recurrent CNVs is similar among children and adolescents with the broader EOP phenotype, this would strongly support the development of guidelines for genomic screening in this population. Such an approach could help aid diagnosis, therapeutic choices, and clinical staging of individuals with EOP, many of whom do not respond to first-line treatments (25). However, there are currently no genomic screening guidelines for children or adults with psychotic disorders (26).

In this study, we aimed 1) to establish the prevalence of recurrent CNVs in our diagnostically heterogeneous EOP cohort (N=137) and compare this prevalence with that found in 5,540 individuals with ASD and 16,504 unselected population control subjects; and 2) to compare the CRS observed among EOP probands to the ASD and control cohorts.

## METHODS

### EOP Samples

Unrelated participants with EOP (N=137) were referred to the Developmental Neuropsychiatry Program at Boston Children's Hospital. Clinical diagnoses were ascertained by a board-certified child psychiatrist (J.G.H.) specializing in EOP. Diagnoses were subsequently confirmed via medical record with a DSM-5 checklist, and a consensus diagnosis was reached (Table 1). Inclusion criteria included having a DSM-5 diagnosis for a current or lifetime axis I psychotic disorder with onset before age 18. Exclusion criteria were substance- or medication-induced psychosis, psychosis secondary to a brain infection (e.g., encephalitis), psychosis due to a neurodegenerative disorder (e.g., Wilson's disease, dystonia with fixed musculoskeletal deformities, Huntington's disease, Friedreich's ataxia, ataxia-telangiectasia, Parkinson's disease), and a severe neurodevelopmental disorder or other impairment affecting ability to describe symptoms or provide other information required for this study. All EOP participants or guardians provided written informed consent (or assent for participants under age 18) on forms approved by the Boston Children's Hospital Institutional Review Board as part of the Manton Center for Orphan Disease Research. After providing consent or assent, each participant provided blood samples.

### ASD and Unselected Populations

We compared the EOP participants to two pooled ASD cohorts: 2,585 children from the Simons Simplex Collection (SSC) (27) and 3,171 probands from the MSSNG database (28). We also compared EOP participants to individuals from three pooled unselected community-based cohorts: IMAGEN (N=1,802) (29), Generation Scotland (N=14,160) (30), and the Lothian Birth Cohort (N=554) (31) (Figure 1). Studies for each cohort were reviewed by local institutional review boards and are described elsewhere (27–31).

### Genotyping and CNV Calling

For the EOP sample, genomic DNA from blood was extracted using standard protocols. Dye-swap array-CGH experiments were performed according to the experimental procedures described by Agilent Technologies (Santa Clara, Calif.) using standard 4×180K Surescan arrays and analyzed with the Agilent Cytogenomics software program. Probe sequences and locations are based on Genome Reference Consortium build 37 (GRCh37/hg19). Three criteria were used to determine the presence of a CNV: 1) at least seven consecutive

**TABLE 1. Diagnostic and demographic breakdown of the early-onset psychosis (EOP) cohort**

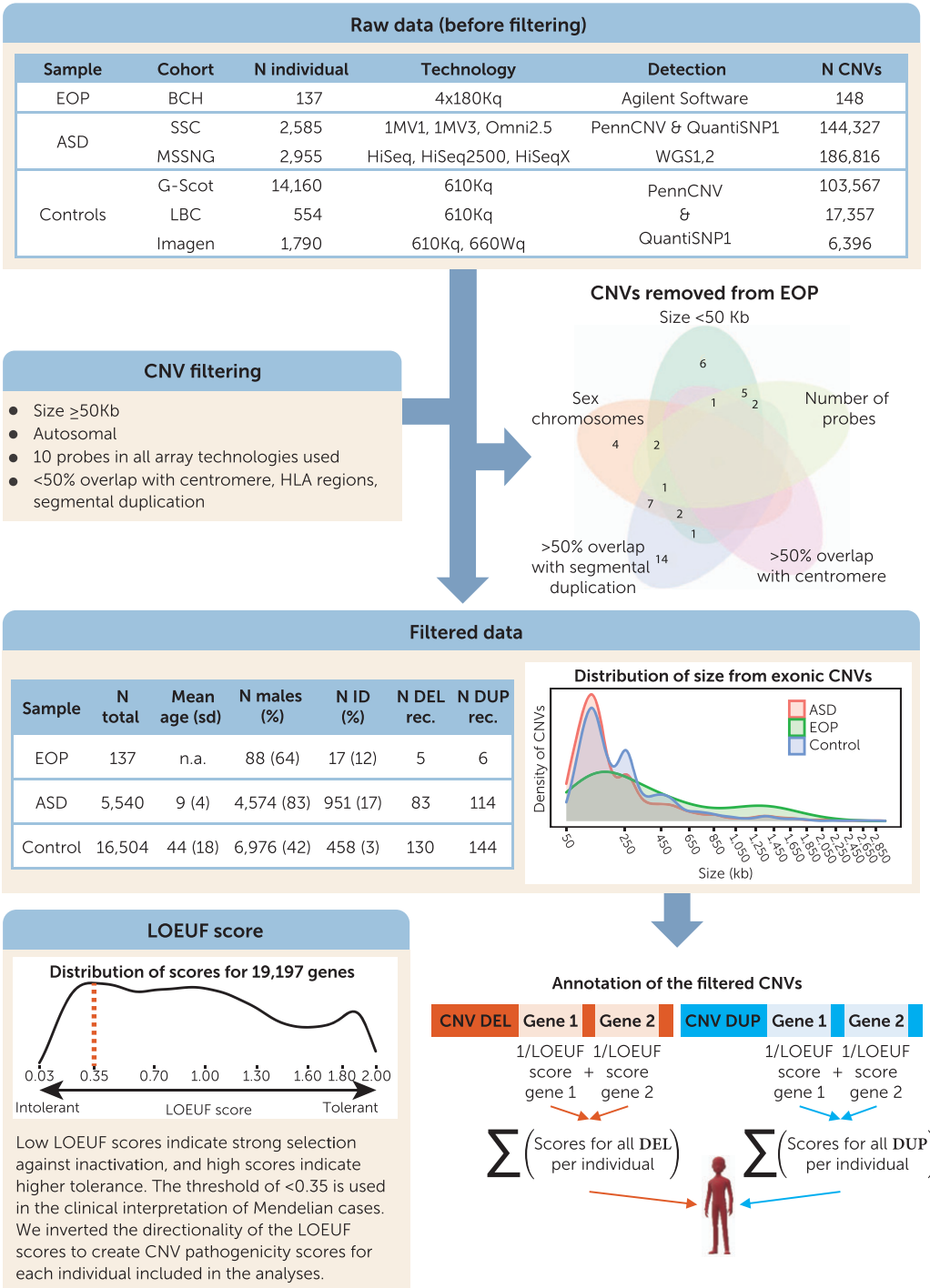
| | Full EOP Sample (N=137) | | EOP Without ASD or Intellectual Disability (N=89) | | EOP Without Schizophrenia (N=99) | | EOP Before Age 13 (N=101) | |
|---|---|---|---|---|---|---|---|---|
| Characteristic | N | % | N | % | N | % | N | % |
| Sex | | | | | | | | |
| Male | 88 | 64.2 | 50 | 56.2 | 63 | 63.6 | 65 | 64.4 |
| Female | 49 | 35.8 | 39 | 43.8 | 36 | 36.4 | 36 | 35.6 |
| Race | | | | | | | | |
| Asian or Pacific Islander | 3 | 2.2 | 3 | 3.4 | 2 | 2.0 | 2 | 2.0 |
| Black/African American | 21 | 15.3 | 15 | 16.9 | 16 | 16.2 | 12 | 11.9 |
| Two or more races | 2 | 1.5 | 2 | 2.2 | 0 | 0.0 | 2 | 2.0 |
| White | 90 | 65.7 | 56 | 62.9 | 65 | 65.7 | 67 | 66.3 |
| Unknown or not available | 21 | 15.3 | 13 | 14.6 | 16 | 16.2 | 18 | 17.8 |
| Ethnicity | | | | | | | | |
| Hispanic/Latino | 22 | 16.1 | 15 | 16.9 | 16 | 16.2 | 15 | 14.9 |
| Not Hispanic/Latino | 91 | 66.4 | 62 | 69.7 | 66 | 66.7 | 67 | 66.3 |
| Unknown or not available | 24 | 17.5 | 12 | 13.5 | 17 | 17.2 | 19 | 18.8 |
| Age at psychosis symptom onset | | | | | | | | |
| <8 years | 38 | 27.7 | 23 | 25.8 | 27 | 27.3 | 38 | 37.6 |
| 8–12 years | 63 | 46.0 | 43 | 48.3 | 43 | 43.4 | 63 | 62.4 |
| 13–18 years | 36 | 26.3 | 23 | 25.8 | 29 | 29.3 | — | — |
| Primary psychosis spectrum disorder | | | | | | | | |
| Schizophrenia | 38 | 27.7 | 21 | 23.6 | — | — | 31 | 30.7 |
| Affective psychosis | | | | | | | | |
| Schizoaffective disorder (bipolar type) | 11 | 8.0 | 6 | 6.7 | 11 | 11.1 | 8 | 7.9 |
| Schizoaffective disorder (depressed type) | 8 | 5.8 | 8 | 9.0 | 8 | 8.1 | 5 | 5.0 |
| Major depressive disorder with psychotic features | 17 | 12.4 | 15 | 16.9 | 17 | 17.2 | 10 | 9.9 |
| Bipolar disorder with psychotic features | 17 | 12.4 | 13 | 14.6 | 17 | 17.2 | 12 | 11.9 |
| Schizophreniform disorder | 2 | 1.5 | 0 | 0.0 | 2 | 2.0 | 2 | 2.0 |
| Brief psychotic disorder | 1 | 0.7 | 0 | 0.0 | 1 | 1.0 | 0 | 0.0 |
| Other specified or unspecified schizophrenia spectrum and other psychotic disorder | 43 | 31.3 | 26 | 29.2 | 43 | 43.4 | 33 | 32.7 |
| Co-occurring diagnoses | | | | | | | | |
| Autism spectrum disorder | 47 | 34.2 | | | | | | |
| Intellectual disability | 17 | 12.4 | — | — | 10 | 10.1 | 11 | 10.9 |
| History of seizures | 24 | 17.5 | 12 | 13.5 | 15 | 15.2 | 19 | 18.8 |

probes in the same direction; 2) 1.5-fold average difference between test and reference DNA; and 3) CNV not present in the within-slide control DNA sample. We applied a previously published pipeline to data from the ASD and control cohorts (16–18). To harmonize the samples, CNVs were filtered by discarding CNVs <50 kb, CNVs that appeared on sex chromosomes, CNVs with >50% overlap with segmental duplication or centromere, and CNVs with <10 probes across all detection technologies used in all included cohorts (see Figure 1).

### Recurrent CNV Analyses

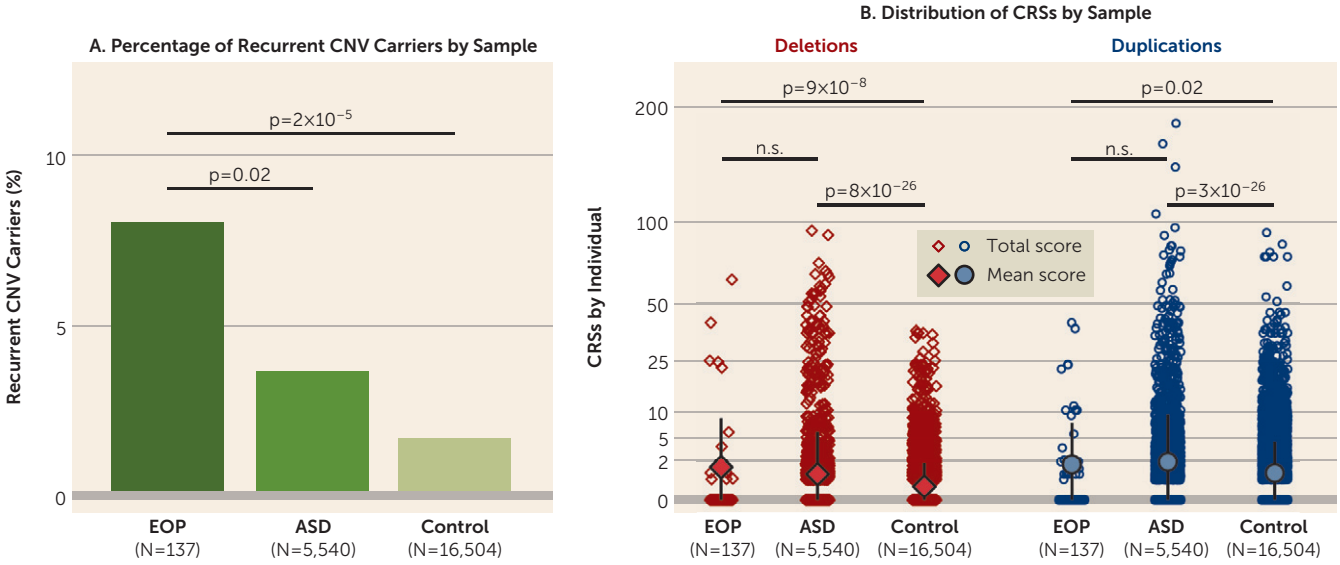
We identified 47 loci and genes previously associated with neurodevelopmental or neuropsychiatric disorders (see Table S1 in the online supplement) to document the frequency of recurrent CNVs in EOP. These loci were defined by >40% overlap with a specific deletion or duplication or if the genes were disrupted by the CNVs (32, 33). Recurrent CNVs share a common size and similar breakpoints, and they recur in multiple individuals in a population (3, 4). Recurrent CNVs generally occur due to nonallelic homologous recombination, which is typically mediated by low-copy repeats (LCRs), resulting in recombination hotspots, gene conversion, and apparent minimal efficient processing segments (5). In contrast, nonrecurrent CNVs are defined as structural variants with dissimilar endpoints or junctions that do not coincide with LCRs but tend to occur in the vicinity of regions that are rich in LCRs, resulting in complex regional genomic architecture. LCRs do not mediate but may stimulate nonrecurrent events (5). Although nonrecurrent CNVs are of different sizes in each patient, they can share a small region of overlap whose change in copy number may result in shared clinical features among different patients (4). Yet, given differences in breakpoints and genes affected, nonrecurrent CNVs, such as private CNVs, are almost impossible to study individually. However, the overall impact of nonrecurrent CNVs can be indexed using genome-wide scores such as the CRS, as described below. To test for differences in the prevalence of recurrent CNVs between the EOP, ASD, and control cohorts, two-sided Fisher's exact tests, computed using the *fisher.test* function in R, were employed. The Benjamin-Hochberg correction for false discovery rate (FDR) was used to adjust for multiple comparisons.

**FIGURE 1. Methodological pipeline for CNV filtering and annotation**[a]




[a] The first table describes CNVs identified in the early-onset psychosis (EOP), autism spectrum disorder (ASD), and unselected control cohorts before filtering. The Venn diagram represents distribution of the EOP CNVs discarded according to the filtering criteria to which they belong: size >50 kb; fewer than 10 probes in at least one of the technologies used; overlap with a centromere or a segmental duplication (<50%); or positioned on a sex chromosome. The second table indicates the prevalence of filtered CNVs and provides demographic and comorbidity data for the EOP, ASD, and control cohorts. The first density plot presents the distribution of the size of the genome-wide CNVs included in the analyses across the different cohorts (EOP in green, ASD in red, unselected control population in blue). The CNV sizes on the x-axis are represented with a square root transformation. The final density plot represents the distribution of LOEUF scores across 19,197 coding genes. A LOEUF score ≤0.35 is the defined clinical threshold for intolerant genes. For CNV annotation, the coding genes totally encompassed in deletions and duplications were identified and the LOEUF score of each gene was attributed. For each individual, the number of genes encompassed and the 1/LOEUF scores for deletions and duplications were summed separately. ASD=autism spectrum disorder; BCH=Boston Children's Hospital; control=unselected control cohort; CNV=copy number variant; EOP=early-onset psychosis; G-Scot=Generation Scotland; ID=intellectual disability; LBC=Lothian Birth Cohort; LOEUF=loss-of-function observed/expected upper bound fraction; N DEL/DUP rec=number of recurrent deletions/duplications previously associated with neurodevelopmental or neuropsychiatric disorders; SSC=Simons Simplex Collection.

**FIGURE 2. CNV burden in the early-onset psychosis, autism spectrum disorder, and unselected control cohorts[a]**




[a] Panel A shows rates of disease-related recurrent copy number variants (CNVs) in individuals with early-onset psychosis (EOP), autism spectrum disorder (ASD) probands, and control subjects from the unselected population. Panel B shows genome-wide gene dosage effects (CNV risk score [CRS]). Scores for deletions and duplications by individual are represented by red diamonds and blue circles, respectively. Individuals without a CNV or with a noncoding CNV have a score of 0. Coding CNVs have scores ranging from 0.5 to approximately 180. The y-axis shows CRS value (root squared of the sum of 1/LOEUF of all genes encompassed in CNVs identified in each individual). The largest diamonds/circles and error bars represent the mean and standard deviation of each group. EOP=early-onset psychosis; ASD=autism spectrum disorder; control=unselected control cohort; DEL=deletions; DUP=duplications; n.s.=nonsignificant.

### CRS: Genome-Wide Dosage Sensitivity Analyses

The CNV risk score (CRS) reflects the probability of intolerance to haploinsufficiency of each gene in the genome that is encapsulated in every CNV identified in an individual, regardless of the population prevalence of the mutation (16–18). In this study, the CRS was calculated as the sum of 1/LOEUF for deletions and duplications separately using our previously published annotation pipeline (see Figure 1). Briefly, each coding gene with all isoforms fully encompassed in filtered CNVs was identified using the Ensembl map (hg19) (34) and was annotated using the inverse LOEUF (1/LOEUF) score (gnomAD, version 2.1.1) (19), which is available for 19,197 genes and ranges from 0.5 (gene tolerant to haploinsufficiency) to 33.3 (gene intolerant to haploinsufficiency). A score of 0 was assigned to individuals with no coding genes encompassed in any CNV. We tested the genome-wide CNV burden with logistic regression models:

$$\ln(\text{odds}[Y_i = \text{diagnosis}_i]) \sim \beta_0 + \beta_1 \text{CRS}_{\text{del}i} + \beta_2 \text{CRS}_{\text{dup}i} + \beta_3 \text{sex}_i$$

$\beta_0$, $\beta_1$, $\beta_2$, and $\beta_3$ are the vectors of coefficients for fixed effects. The logistic regression models were computed using the *glm* function in R.

### Sensitivity Analyses

We performed sensitivity analyses to investigate the robustness of our main results. Specifically, we ran a series of analyses after excluding EOP participants 1) with co-occurring ASD, 2) with co-occurring intellectual disability, 3) with a diagnosis of schizophrenia, or 4) with illness onset before age 13. Each of these sensitivity analyses used statistical models identical to those described above, with smaller sample sizes (see Table 1), and were designed to refine our appreciation of the CNV burden in the EOP cohort relative to the ASD and unselected cohorts.

## RESULTS

### EOP Cohort

A total of 137 EOP patients (88 [64.2%] male) were included in this study (see Table 1). The mean age at psychosis symptom onset was 9.8 years (range, 4–17 years), and 101 (72.3%) patients had psychosis onset before age 13. Thirty-eight (28%) individuals with EOP had co-occurring ASD, 17 (12%) had intellectual disability, and seven (5%) had both ASD and intellectual disability. Thirty-eight (28%) individuals with EOP met DSM-5 criteria for schizophrenia. Sixty-nine CNVs meeting quality control criteria were identified in 55 individuals from the EOP cohort (see Figure 1).

### Prevalence of Recurrent CNVs in the EOP, ASD, and Unselected Population Cohorts

When focusing on recurrent CNVs previously associated with neurodevelopmental or neuropsychiatric disorders (see Table S1 in the online supplement), we found 11 recurrent CNV carriers in the EOP cohort (8.0% of the sample) (Figure 2A, Table 2; see also Table S2 in the online supplement). In contrast, 193 (3.5%) individuals from the ASD cohort and 273 (1.6%) individuals from the unselected population were recurrent CNV carriers. Thus, the prevalence of recurrent CNVs among

**TABLE 2. Enrichment of recurrent CNVs in the early-onset psychosis cohort relative to the autism spectrum disorder and unselected control cohorts[a]**

| Contrast and Group | CNV Carrier Group | | | | |
|---|---|---|---|---|---|
| ASD contrast | ASD carriers | EOP carriers | Odds ratio | 95% CI | p |
| EOP (N=137) | 193 | 11 | 2.42 | 1.16, 4.57 | 0.02 |
| EOP without ASD (N=90) | 193 | 6 | 1.79 | 0.63, 4.11 | 0.16 |
| EOP without intellectual disability (N=120) | 193 | 7 | 1.72 | 0.67, 3.72 | 0.20 |
| EOP without schizophrenia (N=99) | 193 | 8 | 2.43 | 1.01, 5.10 | 0.02 |
| EOP before age 13 (N=101) | 193 | 6 | 1.75 | 0.62, 4.02 | 0.17 |
| Control contrast | Control carriers | EOP carriers | Odds ratio | 95% CI | p |
| EOP (N=137) | 273 | 11 | 5.19 | 2.50, 9.75 | $2\times10^{-5}$ |
| EOP without ASD (N=90) | 273 | 6 | 3.83 | 1.36, 8.77 | $6\times10^{-3}$ |
| EOP without intellectual disability (N=120) | 273 | 7 | 3.68 | 1.43, 7.94 | $4\times10^{-3}$ |
| EOP without schizophrenia (N=99) | 273 | 8 | 5.22 | 2.17, 10.88 | $3\times10^{-4}$ |
| EOP before age 13 (N=101) | 273 | 6 | 3.75 | 1.33, 8.58 | $7\times10^{-3}$ |

[a] Odds ratios are computed using Fisher's exact test. Sensitivity analysis involved excluding individuals in the EOP cohort who had 1) co-occurring ASD, 2) co-occurring intellectual disability, 3) a diagnosis of schizophrenia, or 4) psychosis onset before age 13. ASD=autism spectrum disorder (N=5,540); control=unselected control cohort (N=16,504); CNV=copy number variant; EOP=early-onset psychosis.

children and adolescents with EOP was double that observed among those with ASD (odds ratio=2.42, 95% CI=1.16, 4.57, p=0.02) and five times the rate in the unselected population (odds ratio=5.19, 95% CI=2.50, 9.75, $p=2\times10^{-5}$) (see Table 2).

Three recurrent CNVs were individually enriched in EOP participants relative to unselected control subjects after FDR correction (see Table S3 in the online supplement): 1q21.1 duplication ($p_{FDR}=6\times10^{-4}$) (see Table S4 in the online supplement for a description of 1q21.1 patients), 16p13.11 deletion ($p_{FDR}=0.01$), and 22q11.2 proximal deletion ($p_{FDR}=0.02$). These same loci were enriched in the EOP cohort relative to the ASD cohort, but none survived FDR correction.

### CRSs in the EOP, ASD, and Unselected Population Cohorts

CRSs were higher in EOP participants relative to the unselected cohort for both deletions ($CRS_{del}$: 1.39 vs. 0.23; odds ratio=1.30, 95% CI=1.26, 1.35, $p=9\times10^{-8}$) and duplications ($CRS_{dup}$: 1.63 vs. 0.94; odds ratio=1.09, 95% CI=1.06, 1.12, p=0.02) (Figure 2B and Table 3). Similar results were obtained when comparing ASD participants to the unselected cohort ($CRS_{del}$: 0.86 vs. 0.23; odds ratio=1.24, 95% CI=1.22, 1.26, $p=8\times10^{-26}$; $CRS_{dup}$: 1.88 vs. 0.94; odds ratio=1.12, 95% CI=1.11, 1.13, $p=3\times10^{-26}$). However, CRSs did not differ significantly between the EOP and ASD cohorts ($CRS_{del}$: 1.39 vs. 0.86; odds ratio=1.03, 95% CI=1.01, 1.06, p=0.33; $CRS_{dup}$: 1.63 vs. 1.88; odds ratio=0.98, 95% CI=0.96, 1.01, p=0.61). No sex differences (49 females, 88 males) were observed for the CRS for deletions (odds ratio=1.01, 95% CI=0.96, 109, p=0.70) or duplications (odds ratio=0.97, 95% CI=0.91, 103, p=0.34).

### Sensitivity Analyses: Prevalence of Recurrent CNVs in EOP Subgroups

As can be seen in Table 2, the frequencies of recurrent CNVs in the various EOP subgroups (excluding ASD, excluding intellectual disability, and excluding schizophrenia) did not differ from those observed in the ASD cohort. However, when EOP participants with later symptom onset were excluded, the subgroup had statistically fewer recurrent CNVs than expected when compared with the ASD cohort (p=0.02) (see Table 2). In contrast, each EOP subgroup had a higher prevalence of recurrent CNVs when compared to the unselected control cohort.

### Sensitivity Analyses: CRS in the EOP Subgroups

As in the full sample, we observed no significant differences in dosage sensitivity between any EOP subgroup and the ASD cohort for deletions or duplications (see Table 3). In contrast, the CRS index was elevated for deletions in every EOP subgroup relative to the unselected control cohort. For duplications, the CRS was significantly increased in the EOP subgroups when excluding individuals with intellectual disability (p=0.03) and those with schizophrenia (p=0.02) (see Table 3).

## DISCUSSION

The prevalence of recurrent CNVs in children and adolescents with various EOP diagnoses was far higher than in unselected population control subjects. In contrast, the prevalence of recurrent CNVs in the EOP sample was similar to that in a large ASD cohort. The prevalence of recurrent CNVs in our EOP cohort was also in line with previous reports of individuals with the more restrictive childhood-onset schizophrenia diagnosis (8), even when individuals with a schizophrenia diagnosis were excluded from our cohort. Initially, we selected and analyzed recurrent CNVs that were previously associated with neurodevelopmental (35) and psychotic illnesses (2, 6) (see Table S1 in the online supplement). However, these recurrent CNVs represent only a fraction of all CNVs observed in the population. Thus, we also tested for group differences in CRS, an index of genome-wide dosage sensitivity for deletions and duplications (16). We found higher $CRS_{del}$ and $CRS_{dup}$ in the EOP sample relative to the control population. In contrast, CRS was comparable in the EOP and ASD cohorts. This general pattern of results held even when individuals in the EOP sample with co-occurring ASD or co-occurring intellectual disability were excluded, suggesting that these co-occurring disorders were not completely responsible for the observed CNV burden. Our results indicate that EOP is associated with a substantial CNV burden, strongly

**TABLE 3. Genome-wide CRS differences in EOP relative to ASD and unselected control cohorts[a]**

| Contrast and Group | Deletions | | | | | Duplications | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| ASD contrast | ASD CRS | EOP CRS | Odds ratio | 95% CI | p | ASD CRS | EOP CRS | Odds ratio | 95% CI | p |
| EOP (N=137) | 0.86 | 1.39 | 1.03 | 1.01, 1.06 | 0.33 | 1.88 | 1.63 | 0.98 | 0.96, 1.01 | 0.61 |
| EOP without ASD (N=90) | 0.86 | 1.00 | 1.01 | 0.97, 1.04 | 0.89 | 1.88 | 1.42 | 0.96 | 0.93, 1.00 | 0.48 |
| EOP without intellectual disability (N=120) | 0.86 | 0.53 | 0.94 | 0.89, 0.99 | 0.42 | 1.88 | 1.63 | 0.98 | 0.95, 1.01 | 0.64 |
| EOP without schizophrenia (N=99) | 0.86 | 1.59 | 1.04 | 1.02, 1.07 | 0.24 | 1.88 | 1.76 | 0.99 | 0.96, 1.02 | 0.78 |
| EOP before age 13 (N=101) | 0.86 | 1.25 | 1.03 | 1.00, 1.06 | 0.56 | 1.88 | 1.34 | 0.95 | 0.92, 0.99 | 0.42 |
| Control contrast | Control CRS | EOP CRS | Odds ratio | 95% CI | p | Control CRS | EOP CRS | Odds ratio | 95% CI | p |
| EOP (N=137) | 0.23 | 1.39 | 1.30 | 1.26, 1.35 | $9\times10^{-8}$ | 0.94 | 1.63 | 1.09 | 1.06, 1.12 | 0.02 |
| EOP without ASD (N=90) | 0.23 | 1.00 | 1.24 | 1.19, 1.30 | $3\times10^{-4}$ | 0.94 | 1.42 | 1.07 | 1.04, 1.11 | 0.17 |
| EOP without intellectual disability (N=120) | 0.23 | 0.53 | 1.19 | 1.12, 1.27 | 0.04 | 0.94 | 1.63 | 1.09 | 1.06, 1.12 | 0.03 |
| EOP without schizophrenia (N=99) | 0.23 | 1.59 | 1.31 | 1.27, 1.36 | $3\times10^{-7}$ | 0.94 | 1.76 | 1.10 | 1.07, 1.13 | 0.02 |
| EOP before age 13 (N=101) | 0.23 | 1.25 | 1.34 | 1.28, 1.39 | $4\times10^{-7}$ | 0.94 | 1.34 | 1.06 | 1.03, 1.10 | 0.25 |

[a] The table summarizes the effect of gene dosage on EOP risk. Odds ratios are computed using logistic regressions including the CNV risk score (CRS; sum of 1/LOEUF score) for genes totally encompassed in deletions and duplications as the two main explanatory variables. Odds ratio represents the mean risk conferred by a deletion or a duplication including one intolerant gene (a LOEUF ≤0.35). All models were adjusted for sex. CRSs presented in the table are the mean score by sample. Sensitivity analysis involved excluding individuals in the EOP sample who had 1) co-occurring ASD, 2) co-occurring intellectual disability, 3) a diagnosis of schizophrenia, or 4) psychosis onset before age 13. ASD=autism spectrum disorder (N=5,540); control=unselected control cohort (N=1,650); EOP=early-onset psychosis; LOEUF=loss-of-function observed/expected upper fraction.

suggesting that systematic genetic screening in EOP is clinically warranted.

Given the success of genetic screening in ASD (36), our findings suggest that all children and adolescents with a psychosis diagnosis could substantially benefit from chromosomal microarray testing, with the potential for further testing contingent upon family history and/or clinical features. Universal genetic screening (26) could help disentangle the clinical heterogeneity among youths with EOP (12), potentially leading to specific treatment regimens. As detailed by Moreno-De-Luca et al. (37), genetic diagnoses allow clinicians to communicate more effectively with patients and families, and facilitate genetic counseling. Genetic information could also connect families to additional resources and networks, such as other families with the same CNV. Forming cohorts of patients with the same CNV has yielded valuable information about comorbidities, the range of possible phenotypes, and disease progression in other areas of medicine (38–40). Furthermore, children and adolescents with EOP who carry recurrent CNVs associated with serious nonpsychiatric medical conditions (e.g., cardiovascular abnormalities in 22q11.2 deletion syndrome or the high incidence of hypotonia and epilepsy in 15q11.2 duplication carriers) could be more carefully monitored. Finally, information derived from genetic screening is often invaluable to families of children with ASD (36), helping parents appreciate the biological nature of the illness. Similar genetic information would no doubt be well received by families of children with EOP too. Overall, genetic screening in EOP has the potential to bring us one step closer to true precision medicine in pediatric psychiatry.

Among the recurrent CNVs previously associated with neurodevelopmental and neuropsychic illness, we documented an enrichment for three mutations: 22q11.2 proximal deletion (due to a lack of carriers in the population control), 16p13.11 deletion, and 1q21.1 duplication. Each of these CNVs was reported in the previous childhood-onset schizophrenia study (8) and in ASD cohorts (18). These specific CNVs could be particularly informative to the pathobiology of psychosis and neurodevelopment, particularly since these mutations were commonly observed in a large sample of individuals with idiopathic adult-onset schizophrenia (2). Since the number of CNVs observed is directly related to sample size, it is likely that with a larger EOP cohort, evidence for additional recurrent CNVs will emerge. Thus, it is difficult to speculate about the genetic architecture of EOP and whether child and adolescent psychosis is more strongly influenced by an enumerable set of rare mutations of large effect or by countless common mutations of small effect (e.g., polygenic model), currently the favored model for idiopathic adult-onset psychosis. Further investigations with much larger EOP samples are needed.

Since recurrent CNVs reflect only a fraction of deletions and duplications observed in neurodevelopmental and neuropsychiatric illnesses, we also used the CRS to index dosage sensitivity across the genome. We demonstrated a higher overall burden of genes intolerant to mutation in EOP compared to unselected samples. Moreover, larger effect sizes were observed for deletions

than for duplications. Interestingly, effect sizes observed for EOP were in the same range as those found when comparing individuals with ASD to unselected populations (18), suggesting a major contribution of haploinsufficiency in both EOP and ASD. Additional work is needed to document the relative strength of our CRS score in EOP compared to adult-onset psychotic disorders, as well as to other neurodevelopmental disorders.

A series of sensitivity analyses determined the association of co-occurring neurodevelopmental disorders, psychosis diagnosis, and symptom onset with CNV burden. These analyses documented few differences between the EOP subgroups and the ASD cohort in terms of the frequencies of recurrent CNVs or with regard to genome-wide dosage sensitivity. In contrast, regardless of the portion of the EOP sample studied, we found that individuals with EOP had increased prevalence of recurrent CNVs and increased CRS for deletions relative to the unselected control cohort, with significant findings for the CRS for duplications for specific subgroups. The sensitivity analyses that excluded co-occurring ASD or co-occurring intellectual disability demonstrated that the CNV burden observed in EOP is not simply due to the presence of other neurodevelopmental disorders in the EOP sample. To assess psychiatric diagnostic specificity, one sensitivity analysis excluded EOP participants with a schizophrenia diagnosis. The prevalence and CRS for this subgroup of individuals with affective and other psychoses remained relatively unchanged, suggesting that the specific diagnosis of schizophrenia did not drive the CNV burden observed in the full EOP cohort. Finally, to compare our findings more directly with those reported in childhood-onset schizophrenia (8), we excluded individuals with EOP whose psychosis symptoms manifested at age 13 or older. However, excluding individuals with later symptom onset did not substantively change the pattern of results. Given that these sensitivity analyses included relatively small sample sizes, our study has limited power to detect group differences, and larger samples are needed to fully address these issues.

Our study has several strengths, such as the use of a unique sample of EOP patients with a range of comorbidities, which is highly representative of children and adolescents with psychotic disorders. Moreover, we were able to compare CNV burden in individuals with EOP to individuals with ASD and general population control subjects. However, our EOP sample was small, which is to be anticipated given the rarity of psychosis in children and adolescents. Future studies with larger samples may reveal additional recurrent CNVs or stronger effects of genome-wide duplications, similar to findings in larger ASD or schizophrenia samples (2, 18). Nonetheless, the high frequency of CNVs in our EOP cohort suggests that routine screening for CNVs should be made available to EOP patients and could have important implications for genetic counseling and patient management. These relatively high penetrance risk alleles are also promising targets for biological research aimed at developing animal and cellular models to identify novel disease mechanisms and drug targets for psychotic disorders.

## AUTHOR AND ARTICLE INFORMATION

Early Psychosis Investigation Center (Brownstein, Mollon, Smith, Hojlo, Knowles, Deaso, D'Angelo, Deo, Glahn, Gonzalez-Heydrich), Division of Genetics and Genomics (Brownstein, Smith, Cabral, Li, Bowen, Rao, Genetti, Agrawal, Beggs, Glahn), Manton Center for Orphan Disease Research (Brownstein, Smith, Cabral, Li, Bowen, Genetti, Agrawal, Beggs, Glahn), Department of Psychiatry and Behavioral Sciences (Mollon, Hojlo, Das, Goldman, Garvey, Carroll, Knowles, Deaso, D'Angelo, Deo, Glahn, Gonzalez-Heydrich), Tommy Fuss Center for Neuropsychiatric Disease Research (Deo, Glahn, Gonzalez-Heydrich), and Division of Newborn Medicine (Agrawal), Boston Children's Hospital, Boston; Department of Pediatrics (Brownstein, Smith, Genetti, Agrawal, Beggs, Deo) and Department of Psychiatry (Mollon, Carroll, Knowles, D'Angelo, Deo, Glahn, Gonzalez-Heydrich), Harvard Medical School, Boston; Department of Pediatrics (Jacquemont) and Department of Neuroscience (Douard, Moreau), Université de Montréal, Montreal; Sainte-Justine Hospital Research Center, Montreal (Douard, Saci, Moreau, Huguet, Jacquemont); Department of Biomedical and Health Informatics (Almasy) and Department of Psychiatry (Alexander-Bloch), Children's Hospital of Philadelphia, Philadelphia; Department of Psychiatry, Rutgers–Robert Wood Johnson Medical School, Piscataway, N.J. (Deo); Rutgers University Behavioral Health Care, Piscataway, N.J. (Deo). Lifespan Brain Institute, Children's Hospital of Philadelphia, Philadelphia (Almasy); Department of Genetics, University of Pennsylvania, Philadelphia (Almasy); Department of Pharmacology, Feinberg School of Medicine, Northwestern University, Chicago (Smith).

Send correspondence to Dr. Brownstein (catherine.brownstein@childrens.harvard.edu).

Dr. Brownstein and Ms. Douard share first authorship. Drs. Jacquemont, Glahn, and Gonzalez-Heydrich share last authorship.

This work was funded by the Tommy Fuss Center for Neuropsychiatric Research, Boston Children's Hospital Intellectual and Developmental Disabilities Research Center Molecular Genetics Core Facility supported by P50HD105351 from the NIH Eunice Kennedy Shriver National Institute of Child Health and Human Development, and by NIMH grant U01 MH119690, the Stanley Center of the Broad Institute of MIT and Harvard, and the Jonathan and Robin Klein and Anne and Paul Marcus families.

The authors gratefully acknowledge the patients and families who participated in this study.

Dr. Brownstein has received sponsored research support from NIH, CDC, Alexion Pharmaceuticals, Amgen, Latigo Therapeutics, and Pfizer; she has served a consultant for Deerfield Institute, GLG Inc., and Sanford Health; and she serves on the board of Dream Impact Trust. Dr. Beggs has received sponsored research support from NIH, MDA (USA), Alexion Pharmaceuticals, Audentes Therapeutics, Avidity Biosciences, Chan Zuckerberg Initiative, Dynacure SAS, and Pfizer; he has served as a consultant for Audentes Therapeutics, Biogen, GLG Inc., Guidepoint Global, Hoffman–La Roche, and Kate Therapeutics; and he holds equity in Kate Therapeutics and KineaBio. Dr. Gonzalez-Heydrich holds equity in and is founding head of the scientific advisory board of Mightier/Neuromotion Labs, and he has served as a consultant for Alkermes, Neurocrine, and Sunovion. The other authors report no financial relationships with commercial interests.

Received November 28, 2021; revision received March 27, 2022; accepted April 20, 2022; published online August 24, 2022.

## REFERENCES

1. Girirajan S, Rosenfeld JA, Coe BP, et al: Phenotypic heterogeneity of genomic disorders and rare copy-number variants. New Engl J Med 2012; 367:1321–1331
2. Marshall CR, Howrigan DP, Merico D, et al: Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects. Nat Genet 2017; 49:27–35
3. Malhotra D, Sebat J: CNVs: harbingers of a rare variant revolution in psychiatric genetics. Cell 2012; 148:1223–1241
4. Hastings PJ, Lupski JR, Rosenberg SM, et al: Mechanisms of change in gene copy number. Nat Rev Genet 2009; 10:551–564

5. Gu W, Zhang F, Lupski JR: Mechanisms for human genomic rearrangements. Pathogenetics 2008; 1:4
6. Bergen SE, Ploner A, Howrigan D, et al: Joint contributions of rare copy number variants and common SNPs to risk for schizophrenia. Am J Psychiatry 2019; 176:29–35
7. Kirov G: CNVs in neuropsychiatric disorders. Hum Mol Genet 2015; 24(R1):R45–R49
8. Ahn K, Gotay N, Andersen TM, et al: High rate of disease-related copy number variations in childhood onset schizophrenia. Mol Psychiatry 2014; 19:568–572
9. Fernandez A, Drozd MM, Thümmler S, et al: Childhood-onset schizophrenia: a systematic overview of its genetic heterogeneity from classical studies to the genomic era. Front Genet 2019; 10:1137
10. Nicolson R, Rapoport JL: Childhood-onset schizophrenia: rare but worth studying. Biol Psychiatry 1999; 46:1418–1428
11. Stentebjerg-Olesen M, Pagsberg AK, Fink-Jensen A, et al: Clinical characteristics and predictors of outcome of schizophrenia-spectrum psychosis in children and adolescents: a systematic review. J Child Adolesc Psychopharmacol 2016; 26:410–427
12. Castro-Fornieles J, Baeza I, de la Serna E, et al: Two-year diagnostic stability in early-onset first-episode psychosis. J Child Psychol Psychiatry 2011; 52:1089–1098
13. Patel PK, Leathem LD, Currin DL, et al: Adolescent neurodevelopment and vulnerability to psychosis. Biol Psychiatry 2021; 89:184–193
14. Hubbard L, Rees E, Morris DW, et al: Rare copy number variants are associated with poorer cognition in schizophrenia. Biol Psychiatry 2021; 90:28–34
15. Zarrei M, Burton CL, Engchuan W, et al: A large data resource of genomic copy number variation across neurodevelopmental disorders. NPJ Genom Med 2019; 4:26
16. Huguet G, Schramm C, Douard E, et al: Genome-wide analysis of gene dosage in 24,092 individuals estimates that 10,000 genes modulate cognitive ability. Mol Psychiatry 2021; 26:2663–2676
17. Huguet G, Schramm C, Douard E, et al: Measuring and estimating the effect sizes of copy number variants on general intelligence in community-based samples. JAMA Psychiatry 2018; 75:447–457
18. Douard E, Zeribi A, Schramm C, et al: Effect sizes of deletions and duplications on autism risk across the genome. Am J Psychiatry 2021; 178:87–98
19. Karczewski KJ, Francioli LC, Tiao G, et al: The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 2020; 581:434–443
20. Wainberg M, Merico D, Huguet G, et al: Deletion of loss-of-function–intolerant genes and risk of 5 psychiatric disorders. JAMA Psychiatry 2022; 79:78–81
21. Tammimies K, Marshall CR, Walker S, et al: Molecular diagnostic yield of chromosomal microarray analysis and whole-exome sequencing in children with autism spectrum disorder. JAMA 2015; 314:895–903
22. Council on Children With Disabilities, Section on Developmental Behavioral Pediatrics, Bright Futures Steering Committee, et al: Identifying infants and young children with developmental disorders in the medical home: an algorithm for developmental surveillance and screening. Pediatrics 2006; 118:405–420
23. Hyman SL, Levy SE, Myers SM, et al: Identification, evaluation, and management of children with autism spectrum disorder. Pediatrics 2020; 145:e20193447
24. Miller DT, Adam MP, Aradhya S, et al: Consensus statement: Chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. Am J Hum Genet 2010; 86:749–764
25. Sikich L, Frazier JA, McClellan J, et al: Double-blind comparison of first- and second-generation antipsychotics in early-onset schizophrenia and schizo-affective disorder: findings from the Treatment of Early-Onset Schizophrenia Spectrum Disorders (TEOSS) study. Am J Psychiatry 2008; 165:1420–1431
26. Martin CL, Wain KE, Oetjens MT, et al: Identification of neuropsychiatric copy number variants in a health care system population. JAMA Psychiatry 2020; 77:1276–1285
27. Fischbach GD, Lord C: The Simons Simplex Collection: a resource for identification of autism genetic risk factors. Neuron 2010; 68: 192–195
28. C Yuen RK, Merico D, Bookman M, et al: Whole genome sequencing resource identifies 18 new candidate genes for autism spectrum disorder. Nat Neurosci 2017; 20:602–611
29. Schumann G, Loth E, Banaschewski T, et al: The IMAGEN study: reinforcement-related behaviour in normal brain function and psychopathology. Mol Psychiatry 2010; 15:1128–1139
30. Smith BH, Campbell A, Linksted P, et al: Cohort profile: Generation Scotland: Scottish Family Health Study (GS:SFHS): the study, its participants, and their potential for genetic research on health and illness. Int J Epidemiol 2013; 42:689–700
31. Deary IJ, Gow AJ, Pattie A, et al: Cohort profile: the Lothian Birth Cohorts of 1921 and 1936. Int J Epidemiol 2012; 41: 1576–1584
32. Kasak L, Rull K, Sõber S, et al: Copy number variation profile in the placental and parental genomes of recurrent pregnancy loss families. Sci Rep 2017; 7:45327
33. Tsuang DW, Millard SP, Ely B, et al: The effect of algorithms on copy number variant detection. PLoS One 2010; 5:e14456
34. Frankish A, Diekhans M, Ferreira AM, et al: GENCODE reference annotation for the human and mouse genomes. Nucleic Acids Res 2019; 47:D766–D773
35. Kendall KM, Bracher-Smith M, Fitzpatrick H, et al: Cognitive performance and functional outcomes of carriers of pathogenic copy number variants: analysis of the UK Biobank. Br J Psychiatry 2019; 214:297–304
36. Jeste SS, Geschwind DH: Disentangling the heterogeneity of autism spectrum disorder through genetic findings. Nat Rev Neurol 2014; 10:74–81
37. Moreno-De-Luca D, Ross ME, Ross DA: Leveraging the power of genetics to bring precision medicine to psychiatry: too little of a good thing? Biol Psychiatry 2018; 83:e45–e46
38. Sanders SJ, Ercan-Sencicek AG, Hus V, et al: Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. Neuron 2011; 70: 863–885
39. Ahmad A, Bhattacharya S, Sridhar A, et al: Recurrent copy number variants associated with bronchopulmonary dysplasia. Pediatr Res 2016; 79:940–945
40. Glassford MR, Purcell RH, Pass S, et al: Caregiver perspectives on a child's diagnosis of 3q29 deletion: "We can't just wish this thing away". J Dev Behav Pediatr 2022; 43:e94–e102