



Published in final edited form as:

Am J Psychiatry. 2019 January 01; 176(1): 29–35. doi:10.1176/appi.ajp.2018.17040467.

Joint contributions of rare CNVs and common SNPs to risk for schizophrenia

Sarah E. Bergen, PhD^{1,2}, Alexander Ploner, PhD¹, Daniel Howrigan, PhD³ on behalf of CNV Analysis Group and the Schizophrenia Working Group of the Psychiatric Genomics Consortium, Michael C. O'Donovan, MD, PhD⁴, Jordan W. Smoller, MD, ScD⁵, Patrick F. Sullivan, MD, FRANZCP^{1,6}, Jonathan Sebat, PhD⁷, Benjamin Neale, PhD³, and Kenneth S. Kendler, MD⁸

¹Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

²Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge MA, USA

³Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA, USA

⁴MRC Centre for Neuropsychiatric Genetics and Genomics, Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff, UK

⁵Harvard Medical School, Department of Psychiatry, Massachusetts General Hospital, Boston, MA 02114, USA; Psychiatric and Neurodevelopmental Genetics Unit, Center for Human Genetics Research, Massachusetts General Hospital, Boston, MA 02114, USA

⁶Departments of Genetics and Psychiatry, University of North Carolina, Chapel Hill, NC, USA

⁷Departments of Psychiatry and Cellular and Molecular Medicine, University of California, San Diego, CA, USA

⁸Virginia Institute for Psychiatric and Behavioral Genetics, Department of Psychiatry, Virginia Commonwealth University, Richmond, VA, USA

Abstract

Objective: Both rare copy number variants (CNVs) and common single nucleotide polymorphisms (SNPs) contribute to liability to schizophrenia, but their etiological relationship has not been fully elucidated. We evaluated an additive model, whereby risk of schizophrenia requires less contribution from common SNPs in the presence of a rare CNV and test for interactions.

Method: Genetic data from 21,094 schizophrenia cases and 20,227 controls from the Psychiatric Genomics Consortium were used. We assessed three classes of rare CNVs: those previously associated with schizophrenia, large deletions > 500kb, and total CNV burden. We compared mean

polygenic risk scores (PRS) between subjects with and without rare CNVs and modeled the joint effects of PRS and CNVs on schizophrenia liability using logistic regression.

Results: Schizophrenia cases carrying risk CNVs have lower polygenic risk than those without CNVs, but still higher than controls. For cases carrying known risk CNVs, the PRS was diminished proportional to the effect size of the CNV. The strongly-associated 22q11.2 deletion required little added PRS to produce schizophrenia. Large deletions and increased CNV burden were also associated with lower polygenic risk in cases but not in controls or after removal of known risk CNV carriers.

Conclusions: We found evidence for interactive effects of PRS and previously associated CNVs for risk for schizophrenia, while our results for large deletions and total CNV burden support an additive model. These findings offer insights into the genetic architecture of schizophrenia by illuminating how different established genetic risk factors act and interact to influence liability to schizophrenia.

Keywords

genome-wide association study; schizophrenia; risk profile score; copy number variation; structural variation

Schizophrenia is a serious, chronic mental illness with high heritability (64–81%).¹² Important progress has been made in understanding the genetic basis of schizophrenia. Genome-wide association studies (GWAS) have identified more than 108 single nucleotide polymorphisms (SNPs) which contribute to increased likelihood of schizophrenia³. However, the majority of SNPs contributing to schizophrenia liability fall short of genome-wide significance, and indices of polygenic risk incorporating larger proportions of SNPs have consistently demonstrated highly significant case-control differences^{3,4}.

While common SNPs have weak individual effects (odds ratios, $OR < 1.2$), several rare copy number variants (CNVs) have been identified which impact much more strongly on risk ($ORs = 2–57$)⁵. Furthermore, an increased liability to schizophrenia has been associated with large deletions throughout the genome⁶⁷, and an elevated overall CNV burden^{7, 89}. For the specific deletions and duplications that confer risk for schizophrenia, only 1.4–2.5% of cases carry one of them⁵. Risk for schizophrenia conferred by these CNVs is not deterministic, and many carriers do not develop schizophrenia. It is not known whether the additional factors impacting disease liability are environmental, or reflect genetic variation within the CNV region, or risk variants elsewhere in the genome.

Despite the strong effects from individual CNVs, the aggregate effect of common SNPs are at least an order of magnitude greater⁷¹⁰. Some overlap between GWAS and CNV findings for schizophrenia have been reported^{3, 7}, and cases with associated CNVs still had elevated liability from common SNPs¹¹. However, these two categories of genetic risk have generally been examined separately, and the relationship between them remains poorly understood.

This study investigates the ways in which common SNPs and rare CNVs jointly contribute to risk for schizophrenia. We test a liability threshold model in which SNPs and CNVs act additively to confer disease risk. This model predicts that individuals with schizophrenia

who have large effect CNVs will, on average, have a smaller contribution from common SNPs. We also test for interactions between common SNPs and specific CNVs. A second testable prediction from this model is that among controls, those with large effect CNVs will typically have lower polygenic risk than controls without CNVs.

METHODS

Subjects

This study used subjects from the Psychiatric Genomics Consortium schizophrenia study³ with available CNV data. Genome-wide genotype data from 33 independent, European ancestry case-control samples were used (Supplementary Table 1). Each sample collection was approved by relevant ethical review boards. All participants were at least 18 years of age and provided written, informed consent.

CNV data

CNV data were derived from GWAS arrays and processed using a standardized pipeline by the Psychiatric Genomics Consortium-CNV analysis group (full details in reference¹²). Briefly, multiple calling algorithms were applied to raw Illumina or Affymetrix intensity data from each individual. A consensus CNV call dataset was generated by merging data at the sample level. Following merging, arrays with excessive probe variance or guanine-cytosine bias were removed, as were samples with mismatches in sex, ancestry outliers, >7 mb total CNV burden, or chromosomal aneuploidies. We removed samples with low-quality SNP genotyping or who were related to any other subject. The final dataset of rare, high-quality CNVs retained CNVs ≥ 20 kb, ≥ 10 probes, and frequency < 0.01 . CNVs that overlapped $> 50\%$ with regions tagged as copy number polymorphic on any platform were excluded. Only autosomal chromosomes were used to facilitate comparability between sexes.

A total of 41,310 individuals met these criteria (21,088 schizophrenia cases and 20,222 controls).

Risk CNV classes

Three categories of CNV risk were investigated. First, implicated loci were specific CNVs reported as genome-wide significant (Supplementary Table 2). Carriers were defined as having $\geq 50\%$ reciprocal overlap with reported CNVs (subjects with overlap $< 50\%$ were excluded from all analyses involving implicated loci). For NRXN1 deletions, each exon was considered separately. Six subjects carried two implicated CNVs, and the CNV conferring greatest risk was retained for analysis. Second, large CNV deletions (≥ 500 kb) anywhere in the genome were carried by 722 cases and 477 controls. Third, the total CNV burden for each subject was also examined.

Polygenic risk quantification

We generated risk profile scores by weighting each SNP by its log odds ratio in an independent set of GWAS results and applying these weights to SNPs in a second target dataset. Summed across all SNPs, this yields a risk score for each subject. These are a

continuous and normally distributed measure of schizophrenia liability with highly significant differences between cases and controls⁴³.

Risk profile scores were generated via leave-one-out analyses (see supplement for details³). Briefly, low frequency (<10%), low-quality (imputation INFO <0.9), indels and SNPs in the extended major histocompatibility complex region (chr6:25–34 Mb) were excluded. After removing SNPs in high linkage disequilibrium and “clumping” (i.e. discarding variants within 500 kb and with $r^2 > 0.1$ with a more significantly associated SNP). Polygenic scoring was performed using PLINK¹³ for multiple P-value thresholds (5×10^{-8} , 1×10^{-6} , 1×10^{-4} , 0.001, 0.01, 0.05, 0.1, 0.2, 0.5, 1.0), multiplying the logistic regression weighting (i.e., the natural log of the odds ratio) of each variant by the imputation probability for the risk allele in each individual. The resulting values were summed over each individual to provide a whole genome risk profile score for further analysis.

Scores were then normalized to reduce between-cohort variation (see supplemental information; Supplemental Figure 1). The information contained in the normalized scores was concentrated through principal component analysis (see supplemental information). The first principal component (PRS1) explains 69% of the total variability in the scores (Supplementary Figure 2A), was the only component associated with schizophrenia risk (OR = 2.40, Supplementary Figure 2B) and was used to index polygenic risk. This has the advantage of capturing the majority of polygenic risk in a single variable.

Statistical Models and Hypotheses

Intuitively, if the contributions of PRS1 level and CNVs to risk of schizophrenia sum, we expect lower PRS1 levels among cases carrying CNVs compared to non-carrier cases. In the presence of CNV-mediated risk, a lower PRS1 will be sufficient to put subjects over the threshold for schizophrenia. A similar argument holds for controls, where controls with CNVs and high PRS1 will be underrepresented compared to controls without CNVs.

More formally stated, since PRS1 and CNV status are both positively associated with risk of schizophrenia, we hypothesize that an additive liability model with an increasing link function predicts lower PRS1 values for individuals strongly influenced by the presence of a previously-associated CNV, a large deletion, or high total CNV burden. Because this prediction holds for both cases and controls separately, we can test the following core hypotheses: (a) for schizophrenia cases, the mean PRS1 among CNV carriers/subjects with higher total CNV burden is lower than for non-carriers/subjects with lower total CNV burden and (b) for controls, the mean PRS1 among CNV carriers/individuals with higher total CNV burden is lower than for non-carriers/individuals with lower total CNV burden.

Both hypotheses were tested with respect to three CNV measures, though power was limited for testing in controls, due to the rarity of schizophrenia-associated CNV variants. For specific CNVs and large deletions, we test differences in mean PRS1 levels between carriers and noncarriers using a two-sided Welch t-test. For total CNV burden, we regress PRS1 levels on total CNV burden and use a two-sided Wald t-test to test for negative slopes among cases and controls.

Based on our findings, we also fit logistic regression models with schizophrenia status as outcome, PRS1 and CNV status as predictors, and adjustment for site, sex, CNV quality and five ancestry principal components. These models were fit separately for carriers and noncarriers of specific CNVs. By comparing a series of nested models via likelihood ratio tests and measures of model fit, we could quantify the contribution of PRS1 and CNV both individually and jointly, as well as test for non-additive effects in modelling schizophrenia risk (see supplementary information). For models with statistically significant non-additive effects, we report predicted odds ratios to illustrate the pattern of non-additivity.

The threshold for significant test results was set to 5%. No multiple testing correction was applied because the statistical tests are not independent.

RESULTS

Table 1 displays the numbers of case and control subjects carrying each type of risk CNV. In cases, the mean polygenic risk score for CNV carriers is significantly lower than for noncarriers (PRS1=0.70 in carriers vs. 0.97 in non-carriers, $p=0.03$) (Table 2). This relationship is stronger with increasing risk from the specific CNVs. When the CNVs are divided into three groups based on the OR of their association with schizophrenia (OR of 1–5, 5–15 and >15), only the CNVs with ORs of >15 have a significantly lower PRS1 score than the non-carriers (Table 2). The upper panel of Figure 1 summarizes the relationship for individual CNVs: on average, the mean PRS1 value for carriers of an individual CNV decreases with the effect size (OR) of the CNV. For example, we find that while cases with 15q11.2 deletions (OR=2.2)⁵ have mean PRS1 close to what we see in non-carrier cases, cases with 22q11.2 or 3q29 deletions (OR=28.3–∞ and OR=57.7, respectively)⁵ have much lower PRS1 scores (Figure 1, Supplementary Table 3).

For controls, the relationship is unexpectedly reversed, as carriers of larger effect size CNVs have significantly *higher* mean risk scores (Table 2). Statistical significance and effect size is less clearly tied to reported odds ratios in controls than in cases.

Cases with large deletion CNVs have reduced PRS1 compared to non-carrier cases (PRS1=0.77 vs PRS1=0.98, $p=0.02$). However, upon removal of cases carrying CNVs previously implicated to increase risk for schizophrenia, the results became non-significant (PRS1=0.89 vs PRS1=0.98, $p=0.43$). No statistically significant differences are observed for controls (Table 3a).

Increasing total CNV burden is associated with significantly decreased PRS1 among cases (Table 3b: reduction of mean PRS1 by 1.05 for each 10 kb extra CNV, $p=0.0024$), but not in controls (increased mean PRS1 by 0.19, $p=0.65$). If the CNVs previously implicated in schizophrenia risk are removed, the burden of the remaining CNVs were not significantly associated with PRS1 ($p=0.08$).

Model fitting results

For non-carriers of CNVs previously shown to be associated with schizophrenia (Supplementary Table 4a), PRS1, large deletions and total CNVs are individually significant

(Models 1–3). Both large deletions and total burden add significantly in an additive manner to PRS (Models 4+5), with no indication of significant interactions (Models 6+7).

For carriers of these previously associated CNVs, the GRS, the log(OR) of the specific CNV as well as other large deletions have significant predictive power (Supplementary Table 4b, Models 1–3), but not total burden (Model 4). Adding the log(OR) to PRS1 improves the model significantly (Model 5), and there is a significant interaction (that is, a non-additive effect) between PRS1 and log(OR) (Model 7). Similarly, adding large deletions to PRS1 confers significant improvement (Model 6), and a significant interaction between them exists (Model 8). The interaction parameter is negative in both models 7+8, meaning that increasing PRS1 levels have less impact on risk of schizophrenia in carriers of a specific CNV than in non-carriers. However, once PRS, the CNV effect size, and their interaction have been properly accounted for, large deletions confer no improvement (Model 9).

The interaction between PRS1 and effect sizes for individual CNVs (Model 7 in Supplementary Table 4b) is illustrated in Table 4: we report the predicted OR for schizophrenia associated with an increase of PRS1 by one unit, sorted by reported effect size of the individual CNVs (smallest to largest, Supplementary Table 2); we also include the corresponding predicted OR for non-carriers (based on Model 1, Supplementary Table 4a) as reference. Only for carriers of the three CNVs with the lowest reported effect sizes (15q11.2 deletions, 16p13.11 duplications, and 1q21.1 duplications) do we see statistically significant evidence that an increase in polygenic risk score actually increases the risk of schizophrenia ($p < 5.5 \times 10^{-9}$ for all variants); the associated predicted ORs (1.41–1.56) are slightly in excess of the predicted OR for non-carriers (1.40), though not statistically significantly so ($p > 0.11$ for all variants).

Crucially, the results for the interaction model for specific loci are in line with the results of testing the original two hypotheses for this CNV category: because of the smaller contribution of PRS1 to total risk of schizophrenia in carriers of medium to high-risk loci, the model implies that cases who carry a specific CNV with higher reported risk will have lower mean PRS1 than cases who do not.

DISCUSSION

The goal of this report was to clarify how aggregate measures of common risk SNPs and rare CNVs jointly contribute to risk for schizophrenia. Five results were noteworthy. First, as predicted by an additive model, aggregate affected carriers of previously identified CNVs for schizophrenia had significantly lower PRSs than affected non-carriers. Second, when we subdivided these CNVs by effect size, the significant reduction of PRS was only seen for the CNVs with the strongest impact on schizophrenia risk. Third, while all cases with large deletions cumulatively have a significantly lowered PRS, when we eliminated cases with previously implicated CNVs, this effect disappeared. Fourth, total CNV burden in cases were significantly and inversely related to PRS. The effect was entirely the result of deletions while duplications had no effect. Furthermore, as with the large deletions, when we removed cases with known CNVs, this relationship disappears. Finally, our formal modeling revealed an additive relationship between the PRS score and either total CNV burden or

large CNVs, meaning that the risk for schizophrenia was well captured by simply taking the sum of these two kinds of genetic risk. However, when we examined individual CNVs, we found a more complex relationship: increasing PRS levels had less influence on the risk of schizophrenia for carriers of large effect than on carriers of small to moderate effect CNVs.

Our results are congruent with a prior report that specific previously-associated CNVs require a genomic context of liability to result in schizophrenia¹¹, supporting the conclusion that these loci do not represent fully-penetrant Mendelian forms of illness. Our analyses in 41,321 individuals included the 11,428 subjects previously reported, and in addition to establishing more conclusively that carriers of schizophrenia-associated CNVs generally require elevated genomic risk, we tested and confirmed that these forms of genetic risk act in an interactive manner. Also, the current study was able to evaluate individual CNV-PRS relationships – yielding important results for carriers as well as researchers generating disease models involving these CNVs. Furthermore, this study tested relationships between polygenic risk and other well-replicated categories of CNVs causing risk, large deletions throughout the genome and total CNV burden.

For large deletion carriers and high total CNV burden, the lower PRS1 observed within cases was primarily driven by carriers of the implicated CNVs and not mirrored in the results for healthy controls. This may arise from the genomic locations of these CNVs in the two groups, as the cases will more often carry CNVs intersecting regions of genomic risk for schizophrenia. Furthermore, schizophrenia cases comprise only a small portion of one tail of the liability distribution. Therefore, a small elevation in risk from CNVs in controls is not likely to have a detectable impact at most points along the liability curve. These results also suggest that the specific CNVs conferring substantial risk for schizophrenia have likely all been identified.

The only results inconsistent with our original hypotheses was the observation of greater polygenic loading in controls with any previously implicated CNV compared to non-carrier controls. Ages of the controls were not available. Many of them could be young and not through their age at risk. Also, not all samples used screened controls. Some of the controls could have had or later become schizophrenia cases, imparting the observed results. Since both CNV and polygenic risk could drive behavioral characteristics in a similar direction, assortative mating could produce co-aggregation, but not specifically in controls. This would be more likely in carriers of the low effect size CNVs which are more often inherited.

Different patterns of results may exist across diseases, indicating different genetic architectures. However, one investigation in children with attention deficit hyperactivity disorder (ADHD) found a similar pattern. ADHD subjects with large deletions (>500 kb; N=60) had lower ADHD polygenic risk scores than other affected children (N=421)¹⁵. Additional studies in other complex genetic diseases will offer a broader understanding of the range of genetic architectures underlying neuropsychiatric disorders.

Five potential limitations should be considered in the interpretation of these results. First, even in this study involving the largest number of schizophrenia cases to date with CNV data, power to detect effects within the carriers of the specific risk CNVs is limited by their

rarity. Second, rare single nucleotide variants identified through DNA sequencing comprise a third class of genetic variation contributing to schizophrenia risk¹⁰. Such data were not available on the samples analyzed and so were not incorporated into these analyses. Third, copy number polymorphisms with >1% frequency, which are rarely investigated were not examined here. Fourth, while we have tested interactions between CNVs and aggregate SNPs, epistatic interactions may exist between specific CNVs and specific risk SNPs that are beyond the scope of this study. Fifth, SNPs included in the polygenic scoring falling within the previously associated CNV regions could bias analyses of interactions. Since only 399/102,636 SNPs fell in these loci, this is unlikely to influence our results. Finally, carriers of some implicated CNVs (particularly 22q11.2) who do not develop schizophrenia are unlikely to be recruited as control subjects due to medical problems and/or intellectual impairment. This may produce inflated estimates of schizophrenia risk and complicates interpretation of the relationship between CNV effect sizes and genomic risk. We therefore conservatively used the lower bound estimate of the 22q11.2 effect size. It is worth noting that within a sample of 329 carriers of this deletion, those who developed schizophrenia were significantly more likely to have additional CNVs impacting genes relevant to this disorder¹⁶.

Conclusion

Genetic risk from rare CNVs and common SNPs contribute to liability to schizophrenia. Previously-implicated CNVs (individually and in aggregate), large CNV deletions, and total genomic CNV burden were separately compared against aggregated genomic risk from SNPs. Schizophrenia cases carrying risk CNVs have lower polygenic risk than other cases, but still higher than controls. When these risk CNVs were categorized by their effect size on schizophrenia, lower polygenic risk was only seen clearly in those with the largest effect sizes. Our results also support interactions for PRS and the implicated CNVs and an additive model for the other CNV classes. Comprehensive understanding of schizophrenia etiology should incorporate risk measures from different genomic sources, and the integration of sequencing-derived rare variation and environmental influences along with CNVs and common genetic variation will ultimately offer a more complete picture.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Disclosures and acknowledgements:

The generous contributions of the thousands of participants offering blood samples to enable this and other studies are gratefully acknowledged. This work was supported by a grant from the National Institute of Mental Health, U01 MH094421.

REFERENCES

1. Lichtenstein P, Yip BH, Bjork C, Pawitan Y, Cannon TD, Sullivan PF, et al. Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet*. 2009;373(9659):234–9. [PubMed: 19150704]

2. Sullivan PF, Kendler KS, Neale MC. Schizophrenia as a complex trait: evidence from a metaanalysis of twin studies. *Arch Gen Psychiatry*. 2003;60(12):1187–92. [PubMed: 14662550]
3. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature*. 2014;511(7510):421–7. [PubMed: 25056061]
4. International Schizophrenia Consortium Purcell SM, Wray NR Stone JL, Visscher PM O Donovan MC, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*. 2009;460(7256):748–52. [PubMed: 19571811]
5. Rees E, Walters JT, Georgieva L, Isles AR, Chambert KD, Richards AL, et al. Analysis of copy number variations at 15 schizophrenia-associated loci. *Br J Psychiatry*. 2014;204(2):108–14. [PubMed: 24311552]
6. Bergen SE, O’Dushlaine CT, Ripke S, Lee PH, Ruderfer DM, Akterin S, et al. Genome-wide association study in a Swedish population yields support for greater CNV and MHC involvement in schizophrenia compared with bipolar disorder. *Mol Psychiatry*. 2012;17(9):880–6. [PubMed: 22688191]
7. Szatkiewicz JP, O’Dushlaine C, Chen G, Chambert K, Moran JL, Neale BM, et al. Copy number variation in schizophrenia in Sweden. *Mol Psychiatry*. 2014;19(7):762–73. [PubMed: 24776740]
8. International Schizophrenia Consortium. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature*. 2008;455(7210):237–41. [PubMed: 18668038]
9. Walsh T, McClellan JM, McCarthy SE, Addington AM, Pierce SB, Cooper GM, et al. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science*. 2008;320(5875):539–43. [PubMed: 18369103]
10. Purcell SM, Moran JL, Fromer M, Ruderfer D, Solovieff N, Roussos P, et al. A polygenic burden of rare disruptive mutations in schizophrenia. *Nature*. 2014;506(7487):185–90. [PubMed: 24463508]
11. Tansey KE, Rees E, Linden DE, Ripke S, Chambert KD, Moran JL, et al. Common alleles contribute to schizophrenia in CNV carriers. *Mol Psychiatry*. 2015.
12. CNV Analysis Group and the Schizophrenia Working Group of the Psychiatric Genomics Consortium. A contribution of novel CNVs to schizophrenia from a genome-wide study of 41,321 subjects. *Nat Genet*. 2017;49(1):27–35. [PubMed: 27869829]
13. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81(3):559–75. [PubMed: 17701901]
14. Kendler KS, Eaves LJ. Models for the joint effect of genotype and environment on liability to psychiatric illness. *Am J Psychiatry*. 1986;143(3):279–89. [PubMed: 3953861]
15. Martin J, O’Donovan MC, Thapar A, Langley K, Williams N. The relative contribution of common and rare genetic variants to ADHD. *Transl Psychiatry*. 2015;5:e506. [PubMed: 25668434]
16. Bassett AS, Lowther C, Merico D, Costain G, Chow EWC, Van Amelsvoort T, et al. Rare genomewide copy number variation and expression of schizophrenia in 22q11.2 deletion syndrome. *Am J Psychiatry*. 2017; 174(11):1054–1063. [PubMed: 28750581]

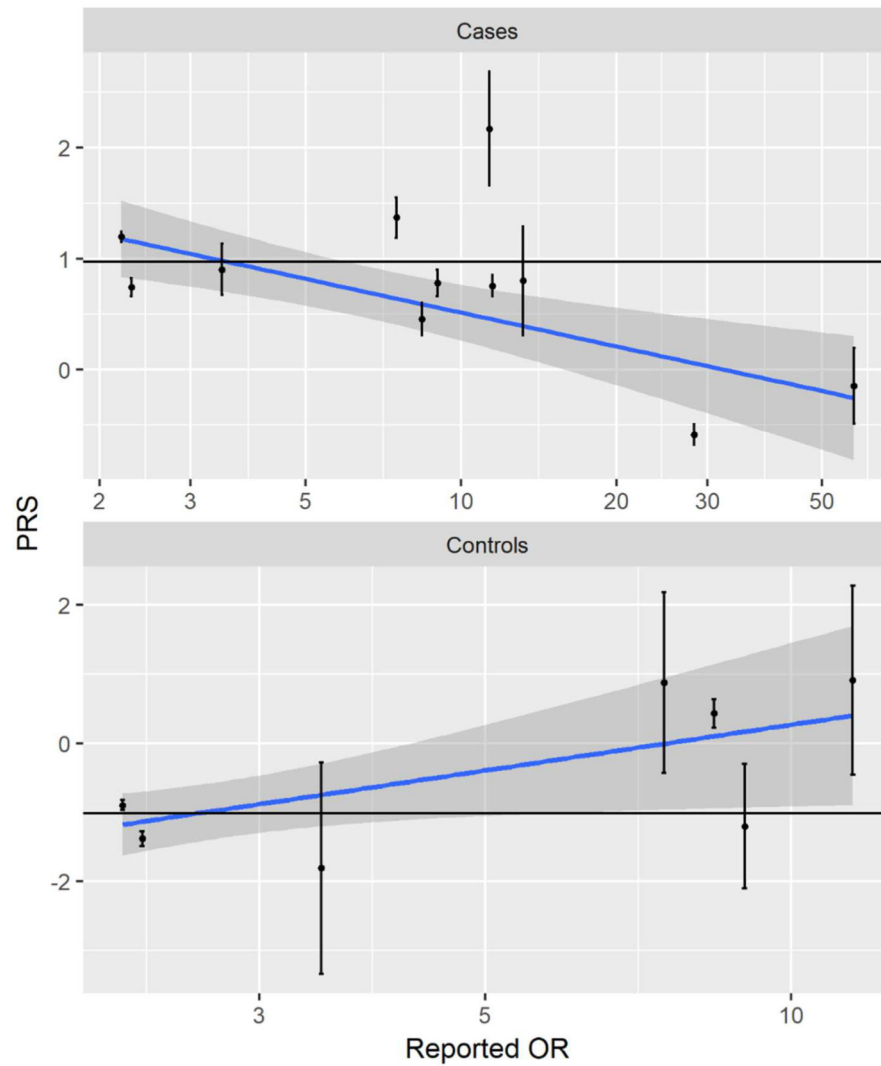


Figure 1: Plot of PRS against reported ORs among carriers of implicated CNVs, summarized by a linear regression line with a 95% confidence band. The solid black shows the average PRS in non-carriers for comparison.

Top: cases, $n = 407$, Pearson $R = -0.18$, $p = 2.5E-4$

Bottom: controls, $n = 115$, Pearson $R = 0.20$, $p = 0.20$

Table 1.

Risk CNV categories and number of carriers by affection status. 21,088 schizophrenia cases and 20,222 controls in total. Individual CNVs are grouped by effect size - low odds ratio (OR; <5), medium OR (5 to 15), and high OR (>15). Dup = duplication. Del = deletion.

CNVs	Case Carriers	Control Carriers
15q11.2 del	97	50
16p13.11 dup	57	46
1q21.1 dup	20	4
15q13.3 del	23	2
1q21.1 del	30	4
NRXN1	36	3
7q11.23 dup	10	0
16p11.2 dup	53	4
15q11.2 dup	11	0
22q11 del	56	0
3q29 del	14	0
Aggregated implicated CNVs	407	115
500 kb deletions	732	479
500 kb deletions (no implicated)	557	441
Mean total kb CNV burden (sd)	312 (492)	279 (415)

Table 2:

Average PRS for carriers and non-carriers, separately for cases and controls

Disease status	CNV status	N	mean PRS	se	p-value
Case	Non-carrier	20640	0.97	0.017	NA
	Carrier	407	0.70	0.120	0.027
	Carrier: OR (1–5)	174	1.01	0.170	0.83
	Carrier: OR (5–15)	163	0.88	0.188	0.64
	Carrier: OR (15–60)	70	–0.50	0.307	5.9e-07
	Control	Non-carrier	20092	–1.01	0.017
Carrier	115	–0.97	0.210	0.85	
	Carrier: OR (1–5)	100	–1.16	0.227	0.547
	Carrier: OR (5–15)	15	0.29	0.436	0.036

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3a:

Average PRS for subjects with and without large scale deletions in all subjects and with carriers of implicated CNVs removed, separately for cases and controls.

Population	Disease status	Large deletion	Count	mean PRS	se	p-value
All subjects	Case	no	20325	0.98	0.017	-
		yes	722	0.76	0.090	0.017
	Control	no	19730	-1.01	0.017	-
		yes	477	-0.92	0.110	0.39
No implicated CNV	Case	no	20093	0.98	0.017	-
		yes	547	0.89	0.102	0.43
	Control	no	19653	-1.01	0.017	-
		yes	439	-0.98	0.115	0.81

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3b:

Slope for regressing PRS on total CNV burden (average change of PRS per 10kb CNV burden) in all subjects and with carriers of implicated CNVs removed, separately for cases and controls and by type of CNV.

Population	Measure	Status	N	Slope	p-value
All subjects	Total Burden	Case	21047	-0.105	0.0024
		Control	20207	0.019	0.6480
	Total Deletions	Case	21047	-0.210	0.0002
		Control	20207	0.078	0.2843
	Total Duplications	Case	21047	-0.043	0.3291
		Control	20207	-0.009	0.8627
No implicated CNV	Total Burden	Case	20640	-0.069	0.080
		Control	20092	0.021	0.618
	Total Deletions	Case	20640	-0.112	0.096
		Control	20092	0.054	0.470
	Total Duplications	Case	20640	-0.047	0.333
		Control	20092	0.006	0.913

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 4:

Odds ratios for a change in +1 PRS, for non-carriers and carriers of implicated CNVs, based on the effect size for CNVs reported in the literature. OR=odds ratios, Lower and Upper=95% confidence limits. Dup = duplication. Del = deletion.

CNV	OR Estimate	Lower	Upper	p-value
Non-carrier	1.40	1.39	1.41	< 2e-16
15q11.2_del	1.56	1.36	1.79	1.6e-10
16p13.11_dup	1.55	1.35	1.77	1.2e-10
1q21.1_dup	1.41	1.26	1.58	5.5e-09
15q13.3_del	1.18	0.99	1.41	0.061
1q21.1_del	1.15	0.95	1.39	0.144
NRXN1	1.13	0.93	1.38	0.216
7q11.23_dup	1.08	0.85	1.35	0.538
16p11.2_dup	1.07	0.85	1.35	0.558
15q11.2_dup	1.04	0.81	1.34	0.767
22q11_del	0.87	0.60	1.27	0.474
3q29_del	0.74	0.46	1.21	0.229

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript