CLINICAL AND COGNITIVE PHENOTYPING OF COPY NUMBER VARIANTS ASSOCIATED WITH NEURODEVELOPMENTAL DISORDERS FROM A MULTI-ANCESTRY BIOBANK

AUTHORS:

Nina Zaks[1], Behrang Mahjani PhD [2,3], Abraham Reichenberg PhD[2,3,4], Rebecca Birnbaum MD[2,5]

AFFILIATIONS:

1. Department of Child and Adolescent Psychiatry, NYU Langone Health
2. Department of Psychiatry, Icahn School of Medicine at Mount Sinai
3. Seaver Autism Center for Research and Treatment, Icahn School of Medicine at Mount Sinai
4. Department of Environmental Medicine and Public Health, Icahn School of Medicine at Mount Sinai
5. Department of Genetic and Genomics Sciences, Icahn School of Medicine at Mount Sinai

Correspondence:
Rebecca Birnbaum, MD
Departments of Psychiatry and Genetics and Genomics Sciences
Icahn School of Medicine at Mount Sinai
1399 Park Avenue, Room 305B
New York, NY 10029
212-585-4657
rebecca.birnbaum@mssm.edu

MANUSCRIPT WORD COUNT: 1,481

KEYWORDS: Copy Number Variants, Neurodevelopmental Disorders, Schizophrenia, Neurocognition, Biobank

ABSTRACT

Clinical biobanks with electronic health records (EHRs) linked to genotype data continue to expand yielding an opportunity to further characterize disease-relevant genomic risk factors, yet few recall-by-genotype studies from biobanks have been published to date. For example, copy number variants (CNVs) that significantly increase risk for multiple neurodevelopmental disorders (NDDs) and negatively affect neurocognition, may present in up to 2% of population cohorts, with public health implications for ascertaining NDD CNV carriers.

From Bio*Me*, a multi-ancestry biobank derived from the Mount Sinai healthcare system (New York, NY), 892 adult participants were recontacted for deep phenotyping, including 335 NDD CNV carriers as well as comparators, 217 individuals with schizophrenia and 340 controls. Clinical and cognitive assessments were administered to each participant. There was no disclosure of genetic information.

Eight percent of recontacted biobank participants completed the study (30 NDD CNV carriers across 15 unique loci, 20 schizophrenia and 23 controls). The study sample had a mean age of 48.8 (10.2) years, was 66% female and of diverse ancestry, 36% African, 34% Hispanic, and 26% European. Overall, 70% of 30 NDD-CNV carriers harbored at least one neuropsychiatric or developmental phenotype, including 40% with mood or anxiety disorders. Further, 22 NDD CNV carriers were significantly impaired compared to controls on digit span backwards (Beta=-1.76, FDR=0.04) and digit span sequencing (Beta=-2.01, FDR=0.04), but higher performing than schizophrenia on verbal learning (Beta=4.5, FDR=0.05).

Thirty NDD CNV carriers were successfully recruited from a multi-ancestry biobank, as well as healthy controls and low-functioning individuals with schizophrenia. Deep phenotyping corroborated past reports, while also identifying discordance with EHRs. Future recall-by-genotype studies may further benchmark the study design and elucidate feasibility.

INTRODUCTION

Clinical biobanks with electronic health records (EHRs) linked to genotype data, continue to proliferate and expand, with potential to identify and clinically characterize disease-relevant genomic risk factors.[1] While numerous *in-silico* biobank analyses have been published as well as biobank-relevant computational methods, relatively few 'recall-by-genotype' studies from biobanks have been published to date.[2, 3] For example, rare copy number variants (CNVs), deletions or duplications greater than one kilobase, are genomic risk factors of high effect for multiple neurodevelopmental disorders (NDDs).[4-6] NDD CNVs also effect neurocognition, with neurocognitive impairment across cognitive domains reported in clinically-ascertained cohorts and healthy controls.[4, 7-10] The prevalence of NDD CNVs across loci may be up to 2% of population cohorts, therefore ascertaining individuals harboring NDD CNVs, even in populations unselected for neuropsychiatric disorders, could have important public health implications.[10, 11]

In a previous study within the Geisinger *MyCode* Community Health Initiative, carriers of 31 neuropsychiatric CNVs were identified and positive reactions to genetic disclosure for 141 carriers of 9 CNVs were reported, however, the cohort was predominantly European, within a relatively contained healthcare system, and without analyses of cognitive function.[11] A recent analysis of Bio*Me*, a multi-ancestry biobank derived from the Mount Sinai healthcare system, with genotype data linked to electronic health records (EHR) of primarily adults, reported 2.5% prevalence of 64 NDD CNVs among approximately 25,000 participants. For NDD CNV carriers, there was an enrichment for congenital disorders and major depressive disorder, as well as an association with multiple medical comorbidities, including hypertension, obesity and increased body mass index, however only EHR-derived outcomes were analyzed.[12] Therefore, in the current study, NDD CNV carriers were recalled from the multi-ancestry Bio*Me* biobank for deep clinical

and cognitive phenotyping beyond EHR-derived outcomes, and then compared to two groups of individuals without NDD CNVs, idiopathic schizophrenia and neurotypical controls.

MATERIALS AND METHODS

IRB: The Bio*Me* biobank is an EHR-linked biobank of genotyped individuals of diverse ancestry, with recruitment across clinics in the Mount Sinai Health System (New York, NY) since 2007, approved by the Mount Sinai School of Medicine Institutional Review Board (IRB). [12, 13] For this study, the IRB approved Bio*Me* data access, participant recontact and study assessments, without genetic disclosure. The study was conducted from November 12, 2020 until March 23, 2023.

Recruitment: Within Bio*Me*, three groups of adults, ages 18-65 were identified for recontact by mail, email or phone: (i) NDD-CNVs: Harboring at least one NDD CNV[12] (ii) Schizophrenia: At least two International Classification of Diseases (ICD10) codes for schizophrenia and without NDD CNVs. (iii) Neurotypical controls: Without ICD-10 code for neuropsychiatric disorders and without NDD CNVs. Respondents were excluded for unstable, severe medical illness, substance or alcohol use disorder, or illness affecting cognition. Chart review was performed for each participant (by board-certified psychiatrist, RB) (Supplementary Methods).

Clinical and Cognitive Assessments: Due to COVID-19, study assessments were performed remotely (Table 1, Supplementary Methods). A 30-minute clinical assessment included the Mini International Neuropsychiatric Interview (MINI v 7.0.2) structured diagnostic interview for DSM-V, as well as an unstructured medical, psychiatric, and developmental history (administered by one board-certified psychiatrist, RB).[14] A 30-minute neurocognitive assessment (administered by NZ) consisted of seven subtests: Digit Span (i) Forward, (ii) Backward and (iii) Sequencing for

attention and executive function; (iv) Category fluency for processing speed and semantic memory; (v) Mayer-Salovey-Caruso Emotional Intelligence Test for social cognition; and Hopkins Verbal Learning Test – Revised (HVLT-R) (vi) immediate and (vii) delayed recall, for verbal learning and memory.[15, 16] Evidence to support remote administration has been demonstrated for schizophrenia and other populations (Supplementary Methods).[17, 18]

Statistical Analyses: Cognitive raw scores were modeled as outcomes for linear regression and group comparisons with covariates of age, sex and ancestry. All analyses were performed using R 4.0.4. Statistical significance was reported using Benjamini-Hochberg false discovery rate (FDR).

Recruitment (Supplementary Figure 1): From Bio*Me* biobank, 892 participants were recontacted including 335 NDD CNV carriers, 217 individuals with schizophrenia, and 340 controls. There was an initial response rate of 18% (15% of NDD CNV carriers, 20% of schizophrenia, 19% of controls) with a drop out of 8%, as participants could not be reached or expressed lack of interest (6%), reported alcohol or substance use (1%), or medical illness confounding cognition (1%). Further, 2% of recontacted participants started the study but were then excluded due to diagnostic discordance (schizophrenia or control not confirmed during study assessment).

RESULTS

Study Sample, Demographics (Table 2A, Supplementary Table 1)*:* Overall, 8% of recontacted biobank participants completed the study, 73 individuals across three groups, 30 NDD CNV carriers, 20 schizophrenia and 23 controls. The overall mean age was 48.8 (10.2) years, 66% female and diverse self-reported ancestry (36% African, 34% Hispanic, 26% European). Age and ancestry did not differ significantly between groups. NDD CNV carriers were predominantly female (87% female) and schizophrenia mostly male (55% male) (p=0.003). Among 30 NDD CNV

carriers, 16 harbored duplications and 14 deletions, across 15 unique NDD CNV loci (Table 2B, Supplementary Table 2).

Study Sample, Healthcare System Utilization: Overall utilization of the healthcare system did not affect recruitment, as the study sample averaged 125 clinical encounters compared to 133 for others recontacted (t=0.30, p=0.76). However, recent engagement with the healthcare system did affect recruitment, as study participants had a clinical encounter on average 12 weeks prior to recontact, compared to 57 weeks for others recontacted (t=7.1, $p=1.1x10^{-10}$).

Clinical Assessment, Results: The study clinical assessments were found to be highly discordant with EHRs, especially for retrospective developmental history elicited during the study, most of which was unreported in EHRs including history of developmental delay, learning disorders or special education (Table 3). Among 30 NDD CNV carriers, MINI assessments identified no cases of schizophrenia, one case of bipolar disorder with psychosis (TAR duplication), and 40% of NDD CNV carriers with mood or anxiety disorder (including five with major depressive disorder and two with obsessive compulsive disorder).

Neurocognitive Assessment, Results: For cognitive analyses, the NDD CNV group was subset to loci also included in previous UK Biobank reports, by excluding the common 15q13.3 (CHRNA7) duplication, 17p12 deletion, and 2q13 (NPHP1) deletion/duplication, thereby yielding 22 NDD CNV carriers of 11 unique loci.[7, 8] The 22 NDD CNV carriers were significantly impaired compared to controls on digit span backwards (Beta=-1.76, FDR=0.04) and digit span sequencing (Beta=-2.01, FDR=0.04), while performing significantly better than the schizophrenia group on HVLT-R Total Immediate Recall (Beta=4.5, FDR=0.05) (Figure 1, Supplementary Table 3). Ancestry correlated with social cognition ($r^2=0.36$, $p=6.5x10^{-7}$), while sex correlated with social cognition

(r^2=0.12, p=0.003) and verbal learning (r^2=0.11, p=0.004). (Supplementary Figure 2). The performance of each NDD CNV carrier across seven cognitive domains is summarily ranked (Supplementary Table 4, Supplementary Figure 3), with varying cognitive performance observed by locus, for example the 22q11.2 deletion carrier among the worst performing and 1q21.1 duplication, the highest performing.

DISCUSSION

In this study, targeted recruitment of 892 adult participants including 335 NDD CNV carriers from Bio*Me*, a multi-ancestry, healthcare-system derived biobank, yielded an 8% rate of enrollment and study completion, relatively low for clinical research recruitment, albeit greatly variable. Notably, low functioning individuals (i.e. with schizophrenia) were also recalled from the biobank at a rate comparable to other groups recruited, including healthy controls. Of 30 NDD-CNV carriers who completed the study, clinical and cognitive phenotypes were overall consistent with previous reports, though variable expressivity of NDD CNVs is well-known.[4, 7, 11, 19, 20] Seventy percent of 30 NDD CNV carriers harbored at least one neuropsychiatric or developmental phenotype, including 40% with mood or anxiety disorders (none with schizophrenia), 13% with speech or motor developmental delays, 27% with learning disorders, and 13% with a history of special education A high discordance between study derived phenotypes and EHRs highlights a potential need for validating EHR data in biobank research, more broadly. In the current multi-ancestry sample, NDD CNV carriers were impaired compared to healthy controls on tests of executive function and working memory, but less impaired than individuals with schizophrenia on verbal learning and memory, corroborating past reports of relatively larger scale, derived from mostly European cohorts.

This study had several limitations. Clinical assessments elucidated individual phenotypes but were underpowered for statistical associations; cognitive analyses may have been underpowered to detect some between group differences. Recruited NDD CNV carriers were mostly female, furthering a female sex bias in the biobank. Analyses were not stratified by sex or ancestry, but pooled analyses were performed with covariates for sex and ancestry. Remote assessments were conducted due to COVID19 using normative cognitive data from in-person testing. CNV loci vary in pathogenicity; analyses were not stratified by CNV locus and included individuals affected and unaffected by neuropsychiatric phenotypes.

The proof-of-concept recall by genotype pilot study recruited NDD CNV carriers from a multi-ancestry biobank along with low-functioning individuals with schizophrenia and healthy controls, for deep phenotyping beyond EHR outcomes, albeit at a low recruitment rate. Future recall-by-genotype studies from biobanks are needed to elucidate the feasibility and generalizability of this study design.

FUNDING:
The study was supported by K23MH112955 (PI: Birnbaum). Dr. Birnbaum is also supported by R21MH137536. Drs. Mahjani and Reichenberg are supported by the Beatrice and Samuel A. Seaver Foundation.

ACKNOWLEDGEMENTS:
We thank individuals within the Bio*Me* Biobank for their participation. We thank colleagues from the Institute of Personalized Medicine at the Icahn School of Medicine for facilitating recall of the BioMe biobank participants: Amanda Merkelson, Sheryl Cruz and Alanna Gomez. We are grateful Dr. Zhongyang Zhang for his perusal of the manuscript and comments.

AUTHOR CONTRIBUTIONS:
RB conceptualized the study, obtained funding and secured IRB approval for the project. RB reviewed biobank data, and RB and NZ interacted with recalled study participants to collect additional data. AR contributed to study design and oversight. NZ, BM and RB contributed to data analyses. All authors reviewed and approved the final draft for manuscript.

DISCLOSURES:
The authors report no biomedical financial interests or potential conflicts of interest.

REFERENCES

1. Lessons learned from the eMERGE Network: balancing genomics in discovery and practice. HGG Adv. 2021;2(1):100018.
2. Bastarache L, Denny JC, Roden DM. Phenome-Wide Association Studies. JAMA. 2022;327(1):75-6.
3. Denny JC, Ritchie MD, Basford MA, Pulley JM, Bastarache L, Brown-Gentry K, et al. PheWAS: demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations. Bioinformatics. 2010;26(9):1205-10.
4. Mollon J, Almasy L, Jacquemont S, Glahn DC. The contribution of copy number variants to psychiatric symptoms and cognitive ability. Mol Psychiatry. 2023;28(4):1480-93.
5. Marshall CR, Howrigan DP, Merico D, Thiruvahindrapuram B, Wu W, Greer DS, et al. Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects. Nat Genet. 2017;49(1):27-35.
6. Sebat J, Lakshmi B, Malhotra D, Troge J, Lese-Martin C, Walsh T, et al. Strong association of de novo copy number mutations with autism. Science. 2007;316(5823):445-9.
7. Kendall KM, Rees E, Escott-Price V, Einon M, Thomas R, Hewitt J, et al. Cognitive Performance Among Carriers of Pathogenic Copy Number Variants: Analysis of 152,000 UK Biobank Subjects. Biol Psychiatry. 2017;82(2):103-10.
8. Kendall KM, Bracher-Smith M, Fitzpatrick H, Lynham A, Rees E, Escott-Price V, et al. Cognitive performance and functional outcomes of carriers of pathogenic copy number variants: analysis of the UK Biobank. Br J Psychiatry. 2019;214(5):297-304.
9. Molloy CJ, Quigley C, McNicholas Á, Lisanti L, Gallagher L. A review of the cognitive impact of neurodevelopmental and neuropsychiatric associated copy number variants. Transl Psychiatry. 2023;13(1):116.
10. Stefansson H, Meyer-Lindenberg A, Steinberg S, Magnusdottir B, Morgen K, Arnarsdottir S, et al. CNVs conferring risk of autism or schizophrenia affect cognition in controls. Nature. 2014;505(7483):361-6.
11. Martin CL, Wain KE, Oetjens MT, Tolwinski K, Palen E, Hare-Harris A, et al. Identification of Neuropsychiatric Copy Number Variants in a Health Care System Population. JAMA Psychiatry. 2020;77(12):1276-85.
12. Birnbaum R, Mahjani B, Loos RJF, Sharp AJ. Clinical Characterization of Copy Number Variants Associated With Neurodevelopmental Disorders in a Large-scale Multiancestry Biobank. JAMA Psychiatry. 2022;79(3):250-9.
13. Abul-Husn NS, Soper ER, Odgis JA, Cullina S, Bobo D, Moscati A, et al. Exome sequencing reveals a high prevalence of BRCA1 and BRCA2 founder variants in a diverse population-based biobank. Genome Med. 2019;12(1):2.
14. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry. 1998;59 Suppl 20:22-33;quiz 4-57.
15. Wechsler D. Wechsler adult intelligence scale (WAIS-IV). San Antonio, TX2008.
16. Nuechterlein KH, Green MF, Kern RS, Baade LE, Barch DM, Cohen JD, et al. The MATRICS Consensus Cognitive Battery, part 1: test selection, reliability, and validity. Am J Psychiatry. 2008;165(2):203-13.

17. Brown T, Zakzanis K. A review of the reliability of remote neuropsychological assessment. Applied Neuropsychology: Adult. 2023:1-7.
18. [Available from: Telepractice and the WAIS–IV: Pearson; [Available from: https://www.pearsonassessments.com/content/dam/school/global/clinical/us/assets/telepractice/guidance-documents/telepractice-and-the-wais-iv.pdf.
19. Kendall KM, Rees E, Bracher-Smith M, Legge S, Riglin L, Zammit S, et al. Association of Rare Copy Number Variants With Risk of Depression. JAMA Psychiatry. 2019;76(8):818-25.
20. Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. Nat Neurosci. 2015;18(2):199-209.

MAIN LEGENDS

Table 1: Study Assessment, Overview. Clinical and cognitive assessments administered to each study participant.

Table 2: Study Sample (A) Demographics of each of three comparator groups who were recontacted and completed study assessments. (B) CNV loci of 30 NDD CNV carriers who completed study assessments.

Table 3: Clinical Features of NDD CNV carriers. Neuropsychiatric and neurodevelopmental clinical features as assessed, across each of three comparator groups.

Figure 1: Cognitive Performance, by Group. Boxplots of cognitive performance for seven cognitive tests, for 22 NDD CNV carriers, compared to 20 individuals with schizophrenia and 23 controls.

Table 1: Study Assessment, Overview. Clinical and cognitive assessments administered to each study participant.

CLINICAL ASSESSMENT

Mini International Neuropsychiatric Interview (MINI)
Medical History
Psychiatric History (including medications and hospitalizations)
Developmental History (Developmental Milestones, Learning or Developmental Delay)
Social, Occupational and Educational History

COGNITIVE ASSESSMENT

Test	Domain	Battery
Digit Span Forward	Attention	WAIS-IV
Digit Span Backward	Executive function, Verbal working memory	WAIS-IV
Digit Span Sequencing	Working memory maintenance	WAIS-IV
Category fluency	Processing speed	MCCB
Hopkins Verbal Learning Test - Revised (HVLT-R)	Delayed recall	MCCB
Hopkins Verbal Learning Test - Revised (HVLT-R)	Immediate recall	MCCB
MSCEIT-ME	Social Cognition, Emotional intelligence	MCCB

Notes: MCCB=Matrics Consensus Cognitive Battery, WAIS=Wechsler Adult Intelligence Scale

Table 2: Study Sample (A) Demographics of each of three comparator groups who were recontacted and completed study assessments. (B) CNV loci of 30 NDD CNV carriers who completed study assessments.

(a) Demographics

			RECONTACTED FROM BIOBANK			COMPLETED STUDY ASSESSMENTS					
	Biobank Participants (n=24,877)	NDD CNV (n=627)	NDD CNV (n=335)	Schizophrenia (n=217)	Controls (n=340)	Total (n=73)	NDD CNV (n=30)	Schizophrenia (n=20)	Controls (n=23)	Group Comparison (Test statistic)	Group Comparison (P-value)
Mean Age, Biobank Enrollment (SD)	50.5 (17.3)	50.4 (17.0)	39.0 (10.8)	45.4 (9.9)	46.2 (5.6)						
Mean Age at Study Assessment (SD)						48.8 (10.2)	47.3 [9.5]	50.2 [10.1]	49.5 [11.2]	0.55	0.58
Female (%)	14,586 (59%)	389 (62%)	219 (65%)	97 (45%)	202 (59%)	48 (66%)	26 (87%)	8 (40%)	14 (61%)	11.96	0.003
Highest Level of Education							14.5 (3.3)	11.9 (3.04)	15.3 (3.4)	6.24	0.003
Ancestry (%)							-	-	-	4.98	0.55
African	7892 (32%)	168 (27%)	91 (27%)	101 (46%)	101 (30%)	26 (36%)	10 (33%)	9 (45%)	8 (35%)	0.77	0.68
European	5965 (24%)	209 (33%)	89 (27%)	19 (9%)	50 (15%)	20 (26%)	10 (33%)	2 (10%)	7 (30%)	3.73	0.15
Hispanic	8536 (34%)	200 (32%)	105 (31%)	88 (41%)	143 (42%)	25 (34%)	9 (30%)	9 (45%)	7 (30%)	1.42	0.49
Other	2484 (10)	51 (8%)	50 (15%)	15 (7%)	25 (7%)	2 (3%)	1 (3%)	0 (0)	1 (4%)	-	-

-p-values for age, sex, and ancestry are for comparison across the three comparator groups; F-statistic for mean age and chi-square value for sex and ancestry.

-Ancestry is as per self-report.[12]

-Thirty NDD CNV carriers completed the study; Cognitive analyses were subset to 22 NDD CNV carriers to match previously reported UKBB analyses excluding four loci: 15q13.3 (CHRNA7) duplication, 17p12 deletion, and 2q13 (NPHP1) deletion/ duplication).[7, 8]

(b) NDD CNV Carriers, by CNV Locus

NDD CNV	hg19 coordinates	Completed Study (n=30)	Biobank vs Study Sample (p-value)
TAR_dup	chr1:145,394,955-145,807,817	3	2.29E-05
1q21.1del**	chr1:146,527,987-147,394,444	1	0.27
1q21.1dup	chr1:146,527,987-147,394,444	1	0.05
2q13del(NPHP1)	chr2:110,862,716-110,983,948	2	0.70
2q13dup(NPHP1)	chr2:110,862,716-110,983,948	3	0.08
15q11.2del	chr15:22,805,313-23,094,530	5	8.01E-05
15q11.2dup	chr15:22,805,313-23,094,530	4	0.39
15q13.3dup(CHRNA7)	chr15:32,017,070-32,453,068	2	0.23
16p13.11del	chr16:15,511,655-16,293,689	2	3.35E-03
16p13.11dup	chr16:15,511,655-16,293,689	2	0.19
16p11.2distal_del**	chr16:28,823,196-29,046,783	1	0.13
16p11.2del	chr16:29,650,840-30,200,773	1	0.53
17p12del	chr17:14,141,387-15,426,961	1	0.18
22q11.2del**	chr22:19,037,332-21,466,726	1	3.46E-03
22q11.2dup	chr22:19,037,332-21,466,726	1	0.22

Notes: p-value is a test of equal proportions comparing the prevalence of each NDD CNV locus in the subset that completed the study with overall BioMe prevalence. See Supplementary Table 2 for further detail.

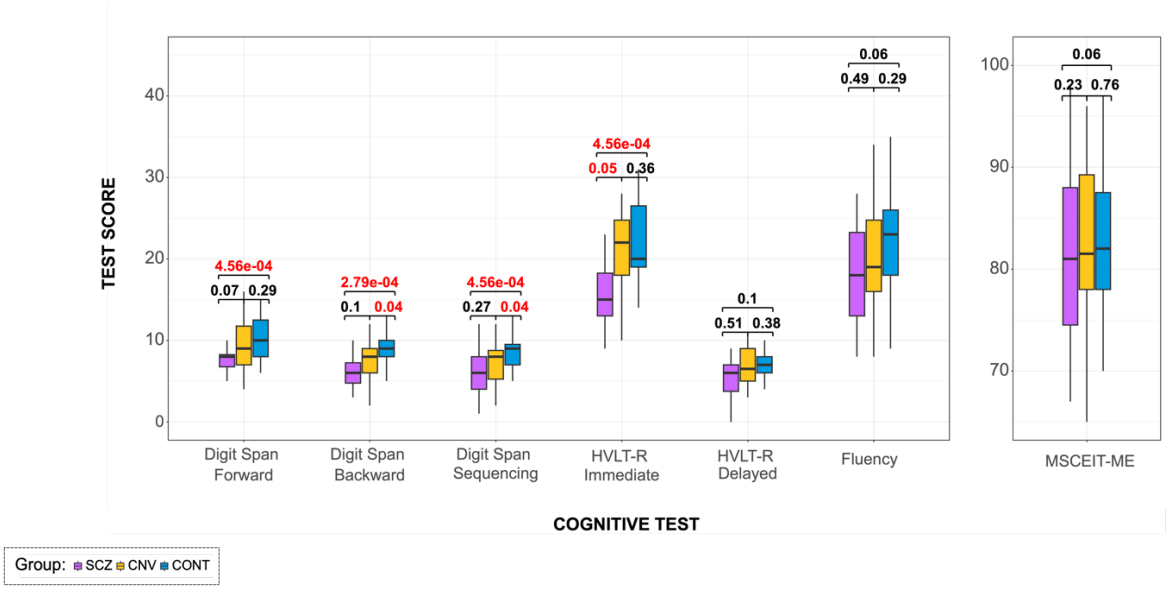
***=NDD CNV loci previously reported to be significantly associated with schizophrenia.*[5]

Table 3: Clinical Features of NDD CNV carriers. Neuropsychiatric and neurodevelopmental clinical features as assessed, across each of three comparator groups.

	NDD CNV (n=30)			SCZ (n=20)		CONT (n=23)	
CLINICAL FEATURE	**N**	**NDD CNV LOCI**	**EHR Unreported/ Discordant**	**N**	**EHR Unreported/ Discordant**	**N**	**EHR Unreported/ Discordant**
MOOD OR ANXIETY DISORDER	12	TAR dup, 1q21.1 del, 1q21.1 dup, 2q13 (NPHP1) dup, 2q13 (NPHP1) dup, 15q11.2 dup, 15q11.2 dup, 15q13.3 dup (CHRNA7), 16p11.2 distal del, 16p13.11 del, 16p13.11 del, 16p13.11 dup	6	0	0	0	0
SEIZURE DISORDER	1	16p11.2 distal del	1	1	1	0	0
DEVELOPMENTAL DELAY (SPEECH OR MOTOR)	4	1q21.1 del, 2q13del(NPHP1), 15q11.2 dup, 16p11.2 del	1	3	1	0	0
CONGENITAL MALFORMATIONS	3	22q11.2 deletion, 2q13 (NPHP1) deletion, 15q11.2 deletion	1	0	0	0	0
LEARNING DISORDER	8	TAR dup; 1q21.1 del; 2q13 (NPHP1) dup, 2q13 (NPHP1) dup, 15q11.2 dup, 15q11.2 dup, 16p11.2 distal del; 22q11.2 del	8	0	0	0	0
HISTORY OF SPECIAL EDUCATION	4	2q13(NPHP1) dup, 15q11.2dup, 1q21.1del, 22q11.2 del	3	10	8	1	1
EXTREME LEARING DIFFICULTIES (WITHOUT LEARING DISORDER DIAGNOSIS)	4	TAR dup; 15q11.2 del; 16p13.11 del; 22q11.2 dup	4	2	2	0	0

Notes: For each group, the count of participants affected for each clinical feature is indicated. For NDD CNV carriers, CNV loci of affected individuals are Indicated; History of Special Education for one control subject was limited to one subject and one semester, without a diagnosed learning disorder. del=Deletion; dup=Duplication

Figure 1: Cognitive Performance, by Group. Neurocognitive performance for 22 NDD CNV carriers, compared to 20 individuals with schizophrenia and 23 controls, across seven cognitive tests



Notes: x-axis is cognitive test, and y-axis is raw cognitive test score. Each cognitive test score was regressed against group status, covarying for age, sex and ancestry. FDR values are indicated for each test. Red indicates statistical significance at FDR threshold of 0.05.