Clinical manifestations of Copy Number Variants among individuals with Autism Spectrum Disorder or Attention Deficit Hyperactivity Disorder: a registry based population study

Maria Rams Vilstrup, Medical Student
Supervisors: Thomas Werge, Professor, Ph.D, Morten Dybdal Krebs, MD, Ph.D Student

Extract from BA project, 2020

Keystrokes (excl. title, figures, tables and references) 17653

Introduction

Autism Spectrum Disorder and Attention Deficit Hyperactivity Disorder are childhood onset disorders with a global prevalence of 1% and 5% respectively[1-2] (male:female ratio 3:1).[3-4] ASD covers a spectrum of conditions with repetitive, rigid behaviours and difficulties in social interactions and communication.[1] ADHD is characterized by deficits of attention, hyperactivity and impulsivity.[5] Comorbid disorders may co-occur for both conditions including anxiety, OCD or intellectual disability (IQ<70), the latter being associated with 45% of ASD-cases.[1]

Both disorders show high heritability (74% for ADHD[6] 64-91% for ASD[7]) and associated genetic variants have been identified,[6-8] including rare pathogenic Copy Number Variants.[9-10] Rare CNVs are genomic deletions or duplications (> 1 kb) occurring with a low population frequency (<1 %),[11] being inherited or arising de novo from mutational events (e.g. non-allelic homologous recombination, responsible for the recurrent CNVs included in this study). CNVs exercise a pleiotropic effect with variable expressivity and penetrance,[12] why CNV carriers may express other phenotypic manifestations including congenital malformations, ID and somatic comorbidities (i.e 22q11 deletion syndrome).[13-14] Several studies have investigated the association between psychiatric disorders and CNVs. However, only few and relatively smaller previous studies have investigated how CNV carriers with a psychiatric diagnosis differ in their phenotypic presentation from other individuals with the same diagnosis.[15-16] To our knowledge, no previous systematic population studies have compared the prevalence of potential predictors of being a CNV carrier and compared these to other individuals with the same diagnosis. Genetic testing in clinical psychiatry is becoming increasingly available[17] and studies have highlighted the beneficial aspects of testing for CNVs to guide clinical treatment and psychoeducation of patients.[18] Currently, the American College of Medical Genetics (ACMG) recommends referral to array-CGH as first tier test for individuals with ASD, ID and/or multiple congenital abnormalities.[19]

The purpose of this study is to address whether features of phenotypic presentation can identify individuals at increased risk of carrying a CNV among individuals diagnosed with ASD or ADHD. The clinical rationale is to help guide decision-making on referral to array-CGH and investigate whether a CNV test has prognostic value. We do this by identifying phenotypic features reported in relation to deletions and duplications at six selected pathogenic CNV loci (1q21, 15q11, 15q13, 16p11, 17q12 and 22q11) from the OMIM database and investigate whether these phenotypic traits can identify individuals at high risk of carrying a CNV among individuals diagnosed with ASD or ADHD in the iPSYCH dataset. Finally, we aim to determine which of the reported phenotypic features are individually associated to CNV-carriers diagnosed with ASD or ADHD in a Danish population cohort.

Material and methods

1-Study population and CNV calling

The present study is a case-only study, comparing CNV-carriers to non-carriers among 17 253 ASD-cases and 20 072 ADHD-cases. Cases are derived from the 2016 updated and expanded version of the iPSYCH Danish case-cohort[20-21] composed of selected Danish singletons born from a known mother from the Danish Civil Registration System between 1. May 1981 and 13. December 2005, alive on their first birthday. Diagnoses of mental disorders are obtained as ICD-10 codes from the Danish Psychiatric Central Research Register before 31. December 2012. 30 000 individuals were randomly sampled from the background population as controls.

DNA from the individuals was extracted from neonatal blood spots, retrieved from the Danish Neonatal Screening Biobank. After genotyping and quality control,[20] the twelve pathogenic CNVs were called using three different algorithms for CNV detection and submitted to visual inspection.[22]

2-Identification of phenotypic features shared between carriers of the twelve pathogenic CNVs

We systematically searched the OMIM database[23] for each of the twelve pathogenic CNVs and sorted out entries marked "*Contiguous gene syndrome*". Phenotypic features were identified from the rubrics "*Clinical features*" and "*Clinical synopsis*". Only features reported across at least two CNVs were retained. When prevalences, odds ratios or p-values related to the features occurred, they were noted *(Supplementary table 1*). OMIM constitutes a compendium of genetic phenotypes based on genetic resources and claims to update daily, why we relied on this database to provide the essential phenotypic features in relation to each CNV.

3-Linking of the phenotypic features reported from OMIM to corresponding variables available in iPSYCH

Phenotypic features reported from OMIM were matched to corresponding variables in the iPSYCH dataset from 2016, as documented in *Supplementary table 2.*

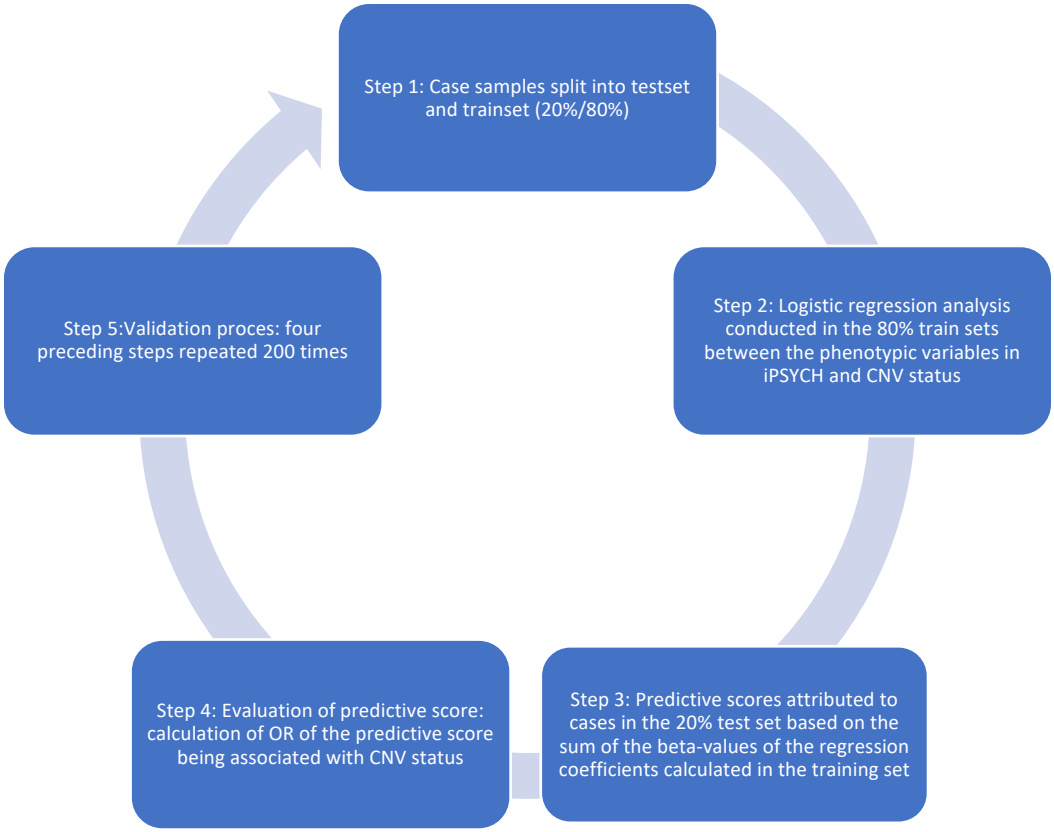
4-Statistical analyses

4-1 Construction of cross-validated prediction model

To investigate whether specific phenotypic traits have predictive value in order to identify CNV carriers, we constructed a predictive scoring model, similar to the Charlson Comorbidity Index. However, we chose to apply the log odds ratios instead of the odds ratios, as discussed to be an improved method in clinical prediction scoring models.[24-25]

The score was constructed through the following steps:

Figure 1: Steps in construction of the predictive score



Step 1:

The two case populations in iPSYCH (ASD and ADHD) were both split randomly into a training (80%) and test set (20%).

Step 2:

In the training set, a univariate logistic regression analysis was conducted between the phenotypic variables in iPSYCH and CNV status of the diagnosed individuals as no correlation was observed between any two of the included phenotypic traits.

Step 3:

In the test set, each individual was attributed a score based on the sum of the beta-values of the regression coefficients calculated in the training set, for the prediction of CNV status of the individual. Only beta-values for significant variables were included in the prediction score, defined as a p-value under 0.05. The attributed scores were scaled according to the population scores.

Step 4:

In order to evaluate the performance of the score, odds ratios for CNV-status being associated with the score were calculated through logistic regression.

Step 5:

The four preceding steps were repeated 200 times to account for empirical variation in the dataset. The obtained OR associated with the predictive score are presented in the histogram of Figure 2 (*cf. Results*).

All analyses were performed in R version 3.6.1.

4-2 Association analysis

A univariate logistic regression analysis was conducted in order to determine which of the previously reported phenotypic features showed true association to CNV carriers in the entire ASD- and ADHD-case groups by mean of the following equation:

$$glm\ (each\ matched\ phenotypic\ variable\ in\ iPSYCH \sim being\ a\ carrier\ of\ any\ of\ the\ 12\ CNVs)$$

A p-value under 0.05 was considered as significant. Whereas parental age was included as a continuous variable in our prediction model, the variable was split into a categorical variable in the association analysis.

Results

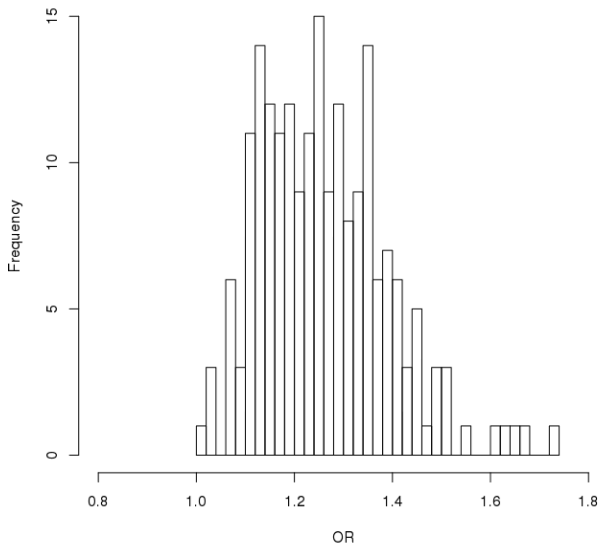
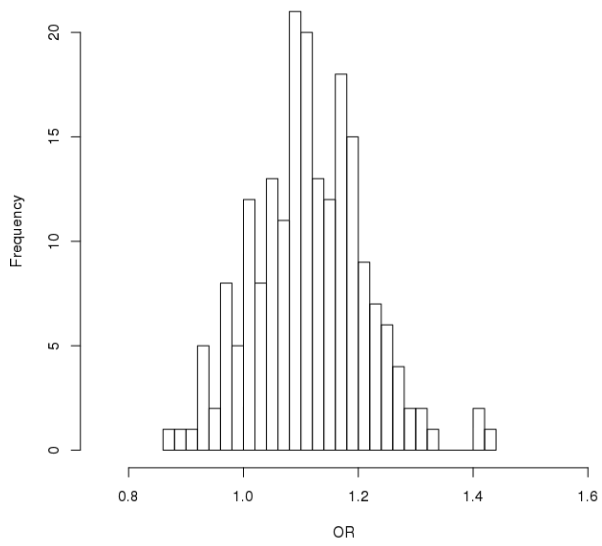
1-Identification of phenotypic features from OMIM and matching to corresponding variables in iPSYCH

We identified 30 phenotypic features related to at least two of the twelve CNVs according to OMIM *(Supplementary table 1*). Among the most frequently reported features were ID (reported in 12/12 CNVs), developmental delay (11/12 CNVs), epilepsy/seizures/abnormal EEG (11/12 CNVs) and skeletal abnormalities or malformations of other organ systems (10/12 CNVS). The features selected from OMIM were matched to 30 corresponding variables in the iPSYCH dataset (*Supplementary table 2).* The number of ASD-only, ADHD-only and comorbid ASD & ADHD cases expressing the phenotypic traits are presented in *Table 1*.

2-Performance of the prediction score

Within the ADHD case group, the prediction score was significantly associated with CNV status (empirical p-value<0.005, mean OR: 1.26). However, the score was not able to significantly distinguish CNV-carriers from non-carriers in the ASD case group (empirical p-value<0.11, mean OR: 1.11). *(Figure 2)*

Figure 2: Odds ratios of the prediction score being associated with CNV status of ASD and ADHD cases after 200 rounds of cross-validation

All ADHD cases				All ASD cases			
Histogram of OR for CNV-status associated with the prediction score							
							
Min OR	*Max OR*	*Mean OR*	*P-value*	*Min OR*	*Max OR*	*Mean OR*	*P-value*
1.018	1.737	1.262	<0.005	0.8764	1.4316	1.1148	<0.113

3- Association analysis

3-1 ASD case group

Comparing Danish ASD-case group to population controls confirmed that the twelve pathogenic CNVs were highly associated with ASD (OR: 1.78, p-value: 1.1. e-13) *(Table 2* in accordance with individual association). [26] Developmental Coordination Disorder (OR: 1.84, p-value: 0.012), congenital malformations (OR: 1.37, p-value: 0.029), other congenital abnormalities (OR: 1.43, p-value: 0.0097) and having a mother or father under 20 years of age was associated with carrying a CNV within the ASD case group (fathers under 20 years: OR: 3.57, p-value: 0.0075, mothers under 20 years: OR: 1.87, p-value: 0.025). A parental history of ID showed association to CNV carriers diagnosed with ASD (OR: 3.90, p-value: 0.0263), whereas ID among cases did not reach significance threshold (OR: 1.29, p-value: 0.059).

3-2 ADHD case group

The association of the twelve CNVs to ADHD as comparison to population controls was replicated in the Danish ADHD-case group[26](OR: 1.74, p-value: 2.16. e-13). ID (OR: 1.97, p-value 1.61 e-07), epilepsy/febrile seizures (OR: 1.44, p-value: 0.019), congenital malformations (OR: 1.37, p-value: 0.029) and other congenital abnormalities (OR: 1.38, p-value: 0.022) were associated to CNV-carriers diagnosed with ADHD. OCD (OR: 0.36, p-value: 0.012), anxiety (OR: 0.61, p-value: 0.015) and mood disorder (OR: 0.68, p-value: 0.028) showed nominally negative association to CNV carriers within the ADHD case group. Parental ID did not reach significance threshold (OR: 2.66, p-value: 0.060).

3-3 Supplementary analyses

A few supplementary analyses were performed in order to achieve better understanding of the association tests (*Table 2)*, especially the discrepancy in features showing association in one diagnostic case group only. We found that the association of ID and epilepsy/febrile seizures to CNV carriers in the ADHD case group was slightly less significant but stronger after excluding comorbid ADHD&ASD cases (ID: OR: 2.22, p-value 4.33 e-07, epilepsy/febrile seizures: OR: 1.45, p-value 0.039) and ID was connected mild ID (OR: 1.96, p-value: 8.48 e-05). Similarly, the association of Developmental Coordination Disorder to CNV carriers within the ASD group was less significant but stronger when excluding comorbid ASD&ADHD cases (OR: 1.97, p-value: 0.016). Finally, the association of younger parental age to CNV carriers found in the ASD-group was replicated in the population control group (paternal age<20 years: OR: 2.69, p-value: 0.032, maternal age<20 years: OR: 1.91, p-value: 0.016).

Discussion and conclusion

Conclusion

We found a broad spectrum of reported phenotypic traits in the OMIM database and were able to link the majority of these to variables for which we have data from national registries. We found that a prediction score based on these variables could significantly separate CNV carrying individuals from other individuals with ADHD, ID appearing to be of particular importance. However, we did not observe this in the ASD case group. This indicates that stratification based on phenotypic characteristics, particularly the presence of ID may help guide the clinical decisions on referral to genetic testing in individuals with ADHD. This does not seem to be straightforward for ASD. Comparing our results to the ACMG guidelines, congenital abnormalities seem to have predictive value among ASD- and ADHD cases, whereas ID was only a significant predictor within the ADHD-case group. Also, our study suggests that additional phenotypic characteristics may help guide clinical decision making on referral to genetic testing.

Limitations

First, the choice of investigating phenotypical features broadly related to twelve CNVs might have been at expense of identifying highly specific traits characterising particular CNVs. However,

our search from OMIM confirmed, that there is a great overlap in features related to CNV carriers at the selected loci. As pathogenic CNVs are rare, our choice was also made in order to draw conclusions with broader clinical utility and gain statistical power to our analyses.

Second, the phenotypic traits reported from OMIM were matched to hospital discharge diagnoses making undiagnosed manifestations of a phenotype impossible to study, e.g. certain behavioural traits or infections not requiring hospitalisation. Also, height and weight, while highly accessible in a clinical setting and possibly relevant predictors of being a CNV carrier, were unavailable in this study.

In addition, the study population is relatively young, hence later onset comorbid disorders such as SZ reported in relation to CNV carriers might have been underrepresented. Likewise, CNV carriers with a particularly severe phenotype might have been excluded from our study population since an inclusion criterion in iPSYCH was to be alive a one year of age. Indeed, congenital malformations at heart/in major veins showed a tendency of association to CNV carriers with ASD but was not significant (OR: 1.72, p-value: 0.086), possibly because severe malformations among CNV carriers might have resulted in an early death (consistent with severe case reports at OMIM).

Discrepancy in performance of our prediction score between the two case groups

Our prediction score was able to significantly identify individuals at increased risk of carrying a CNV within the ADHD case group only. The discrepancy in performance of the score may be explained by the observations from our association test (*Table 2*). Seven out of the 30 included variables in our prediction model were associated to CNV carriers in the ADHD case group including ID (OR: 1.97, p-value: 1.61e-07) vs. five variables in the ASD case group. Hence, the predictive model might have had fewer and weaker variables available when attributing the predictive score to each individual in the test-set within the ASD case group.

Differences and overlaps in association of phenotypic features in the two case groups

Only congenital malformations and other congenital abnormalities were significantly associated with CNV carriers within both case groups. This is in consistency with the frequent reports of malformations at OMIM (10/12 CNVs).

ID and epilepsy/febrile seizures, frequently reported in relation to the twelve CNVs at OMIM, were only associated to CNV carriers diagnosed with ADHD. Surprisingly, we found a nominally significantly negative association between CNV-carrying ADHD-patients and comorbidity with OCD, anxiety and mood disorders, these disorders also being less frequently reported from OMIM (<5/12 CNVs). While other studies suggest that the implication of neuropsychiatric CNVs in Bipolar Disorder is more ambiguous,[27] the negative association with OCD and depression is more surprising because studies have identified an increased burden of neuropsychiatric CNVs for both disorders[28-29]. Our findings could be due to the choice of studying all twelve CNVs as a single group, or to variable expressivity. It is also thinkable, that very severe, early onset disorders, such as ID, may either mask for later onset of other disorders or make clinicians less likely to report such disorders to the registers.

In the ASD case group, ID did not reach significance threshold, but parental ID and DCD were significant predictors of being a CNV carrier. DCD was not directly reported at OMIM, but other studies have identified an increased burden of CNVs among DCD-cases.[30]

Two different phenotypic expressions of CNV carriers within the two respective diagnostic groups or not?

We observed a strengthened association of ID with CNV-carriers in the ADHD-case group and of DCD to CNV-carriers in the ASD case group when excluding comorbid cases, suggesting there might exists truly differing phenotypic features distinguishing CNV carriers within the two diagnostic groups, in spite of the genetic and neurobiological features shared between ASD and ADHD.

However, another interpretation of the differing phenotypic pictures is that comorbidities such as ID are more prevalent among ASD patients (17% of all ASD-only cases presented with ID versus 8.8% of all ADHD-only cases, *Table 1*), and consequently less specific for being carrier of a pathogenic CNV in the ASD case group but possibly the manifestation of other genetic variants.

Finally, one could hypothesize that CNV carriers present with a set of recognizable phenotypic features in a clinical setting, but not all showed statistical significance in both groups in our study, because undiagnosed manifestations of a given trait escaped our analysis (aforementioned limitation).

Parental age

We observed an association of neuropsychiatric CNVs with a younger parental age for the ASD- and control groups, but no associations with an older parental age. A higher burden of de novo CNVs among offspring of younger fathers has been reported by other studies, possibly due to immature spermatids or lower activity in DNA repair mechanisms in younger parents. [16-31] However, the observed association could also reflect a higher burden of inherited CNVs, if CNV carriers tend to have children at a younger age. Since parental genotypes were unavailable in this study, we were not able to make further conclusions about these observations.

While the study provided novel insights as discussed above, further studies are needed to elucidate to what extend the observed differences in phenotypic associations with CNV carriers observed for the two case groups reflect that other factors such as common polygenic variants, rare protein truncating variants and environmental exposures may also affect the phenotypic presentations. It would be interesting to replicate these analyses for other diagnostic case groups, such as schizophrenia, and investigate whether similar observations were to be seen.

Tables

Table 1: Number of individuals with ASD, ADHD or comorbid ASD & ADHD expressing the phenotypic features reported from OMIM

Phenotype	ASD* (n=13168)		ADHD* (n=15987)		ASD & ADHD* (n=4085)	
	n	**%**	**n**	**%**	**n**	**%**
Males	10 025	76.1	11 190	70.0	3300	80.8
CNV carriers	269 (Out of 8838)**	3.0	318 (Out of 10 578)**	3.0	83 (Out of 2898)**	2.9
	Mean (years)	**SD**	**Mean (years)**	**SD**	**Mean (years)**	**SD**
Age of diagnosis	11.1	5.5	13.9	6.7	ASD:11.0 ADHD11.1	ASD: 5.4 ADHD: 5.3
*Table 1 presents the number of exclusive ASD cases, exclusive ADHD cases and comorbid ASD&ADHD cases presenting with the reported phenotypic features from OMIM. **Number of individuals in the case-group for which information about the given phenotypical feature was available. *1)Psychiatric disorders and behavioural psychiatric manifestations*						
	n	**%**	**n**	**%**	**n**	**%**
Schizophrenia or psychotic disorder	1042	7.9	1277	8.0	266	6.5
Mood disorder (bipolar or depression)	1754	13.3	2 598	16.3	386	9.4
Anxiety	1303	9.9	1922	12.0	380	9.3
OCD	710	5.4	644	4.0	206	5.0
Anorexia nervosa	399	3.0	387	2.4	73	1.8
Intellectual disability (IQ<70)	2234 (Out of 13124)**	17.0	1413 (Out of 15972)**	8.8	669 (Out of 4084)**	16.3
Number of individuals diagnosed with ASD or ADHD earlier than mean age of diagnosis	9308	70.7	11515	72.0	2207***	54.0
*** Number of individuals having obtained both diagnoses before mean age of diagnosis of both disorders. *2)Developmental disorders*						
	n	**%**	**n**	**%**	**n**	**%**
Developmental coordination disorder	327 (Out of 13124)**	2.5	313 (Out of 15972)**	2.0	167 (Out of 4084) **	4.1
Developmental delay	1443 (Out of 13124)**	11.0	684 (Out of 15972)**	4.3	421 (Out of 4084)**	10.3
3)Growth and dysmorphic features						
	n	**%**	**n**	**%**	**n**	**%**
Craniofacial dysmorphism	270 (Out of 13124)**	2.1	326 (Out of 15972)**	2.0	71 (Out of 4084)**	1.7
Congenital malformation	593 (Out of 13013)**	4.6	527 (Out of 15839)**	3.3	190 (Out of 4023)**	4.7
Malformation at birth in heart/major veins	247	1.9	270	1.7	83	2.0
Other congenital abnormality	1844 (Out of 13124)**	14.1	1812 (Out of 15972)**	11.3	518 (Out of 4084)**	12.7
Microcephaly	281 (Out of 5286)**	5.3	364 (Out of 5435)**	6.7	127 (Out of 2356)**	5.4
Macrocephaly	168 (Out of 5109)**	3.3	155 (Out of 5314)**	2.9	94 (Out of 2281)**	4.1

Apgar score under 7	120 (Out of 13009)**	0.9	127 (Out of 15801)**	0.8	37 (Out of 4026)**	0.9
Premature birth (under 37 weeks of gestation)	800 (Out of 13020)**	6.1	1112 (Out of 15810)**	7.0	290 (Out of 4061)**	7.1
Small for gestational age (under 10th percentile)	1474 (Out of 12931)**	11.4	2023 (Out of 15712)**	12.9	520 (Out of 4034)**	12.9
4)Somatic comorbidities						
	n	**%**	**n**	**%**	**n**	**%**
Epilepsy/ febrile seizures	1491 (Out of 13125)**	11.4	1317 (Out of 15972)**	8.2	419 (Out of 4084)**	10.3
GI infection	1425	10.8	1898	11.8	469	11.5
Kidney/urinary tract infection	321	2.4	645	4.0	105	2.6
Unspecified infection	3976	30.2	5556	34.7	1271	31.1
Unspecified autoimmune disease	264	2.0	286	1.8	59	1.4
5)Family history and pattern of inheritance						
	n	**%**	**n**	**%**	**n**	**%**
Mother with ID	25 (Out of 13021)**	0.2	38 (Out of 15796)**	0.2	11 (Out of 4028)**	0.3
Father with ID	12 (Out of 13010)**	0.1	24 (Out of 15781)**	0.2	7 (Out of 4023)**	0.2
Mother with epilepsy or febrile seizures	320 (Out of 13021)**	2.5	421 (Out of 15796)**	2.7	140 (Out of 4028)**	3.5
Father with epilepsy or febrile seizures	276 (Out of 13010)**	2.1	421 (Out of 15781)**	2.7	121 (Out of 4023)**	3.0
Mother with any diagnosis of psychiatric disorder	2592 (Out of 13021)**	19.9	4030 (Out of 15796)**	25.5	970 (Out of 4028)**	24.1
Father with any diagnosis of psychiatric disorder	1807 (Out of 13010)**	13.9	3021 (Out of 15781)**	19.1	682 (Out of 4023)**	17.0
	Mean (years)	**SD**	**Mean (years)**	**SD**	**Mean (years)**	**SD**
Maternal age at birth	28.7	5.0	27.2	5.1	28.3	5.2
Paternal age at birth	31.7	6.2	30.1	6.1	31.1	6.2

Table 2: Results of association test for each individual phenotypic feature from OMIM in the ASD and ADHD case groups

I)Association test within group of all individuals with an ASD diagnosis (including cases of comorbid ASD&ADHD)							
Phenotype	**CNV carriers (n=)**		**Non-CNV carriers (n=)**		**OR**	**CI**	**P-value**
	Expressing phenotype (n=)	**Not expressing phenotype (n=)**	**Expressing phenotype (n=)**	**Not expressing phenotype (n=)**			
ASD	352	332*	11384	18745*	1.78	1.53-2.09	1.1e-13
* Comparison of number of CNV carriers in the Danish ASD case group vs. Danish control group							
1)Psychiatric disorders and behavioural psychiatric manifestations							
ADHD	83	269	2815	8569	0.94	0.72-1.20	0.62
SZ or psychotic disorder	26	326	718	10666	1.19	0.77-1.74	0.41
Mood disorder (bipolar or depression)	37	315	1270	10114	0.93	0.65-1.30	0.71
Anxiety	29	323	1048	10336	0.89	0.59-1.28	0.54
OCD	11	341	589	10795	0.95	0.30-1.03	0.090
Intellectual disability (IQ<70)	73	279	1923	9461	1.29	0.98-1.66	0.059
Anorexia nervosa	5	347	266	11118	0.60	0.21-1.32	0.27
	Mean age of diagnosis of all affected individuals (years)		**Mean age of diagnosis of CNV carriers (years)**		**OR****	**CI**	**P-value**
Age of ASD diagnosis	11.1		11.0		0.89	0.70-1.07	0.17
Age of comorbid ID diagnosis	9.7		10.5		0.59	0.37-0.96	0.034
**OR for being a CNV carrier associated with obtaining the diagnosis earlier than the mean age of diagnosis of all ASD-diagnosed individuals							
2)Developmental disorders							
Phenotype	**CNV carriers (n=)**		**Non-CNV carriers (n=)**		**OR**	**CI**	**P-value**
	Expressing phenotype (n=)	**Not expressing phenotype (n=)**	**Expressing phenotype (n=)**	**Not expressing phenotype (n=)**			
Developmental coordination disorder	19	333	342	11042	1.84	1.11-2.88	0.012
Developmental delay	47	305	1243	10141	1.26	0.91-1.70	0.15
3)Growth and dysmorphic features							
Craniofacial dysmorphism	7	345	213	11171	1.06	0.45-2.11	0.87
Congenital malformation	20	329	559	10709	1.37	1.02-1.79	0.029
Malformation at birth in heart/major veins	11	341	210	11174	1.72	0.87-3.03	0.086
Other congenital abnormality	65	287	1550	9834	1.43	1.08-1.88	0.0097
Microcephaly	***	***	***	***	0.33	0.08-0.87	0.056
Macrocephaly	8	148	194	5530	1.54	0.69-2.99	0.24
Apgar 5 score under 7	***	***	***	***	0.61	0.1-1.89	0.33

Premature birth (under 37 weeks of gestation)	21	329	671	10647	1.01	0.63-1.55	0.96
Small for gestational age (under 10th percentile)	49	299	1302	9949	1.25	0.91-1.68	0.15
***Number of individuals not presented, when the group is constituted by four or less than four individuals							
4)Somatic comorbidities							
Epilepsy/febrile seizures	43	309	1239	10145	1.13	0.81-1.56	0.43
GI infection	35	1205	317	10179	0.93	0.64-1.31	0.70
Kidney/urinary tract infection	9	343	268	11116	1.08	0.51-2.01	0.81
Unspecified infection	121	231	3455	7929	0.18	0.96-1.50	0.11
Unspecified autoimmune disease	5	347	203	11181	0.79	0.28-1.74	0.62
5)Family history and pattern of inheritance							
Mother or father with ID	***	***	***	***	3.90	0.93-11.2	0.026
Mother or father with epilepsy or febrile seizures	18	329	562	10670	1.04	0.62-1.63	0.88
Maternal age under 20 years	14	338	246	11138	1.87	1.03-3.13	0.025
Maternal age under 25 years	87	265	2302	9082	1.30	1.01-1.65	0.040
Maternal age under 30 years	226	126	6422	4962	1.36	1.11-1.73	0.0038
Maternal age above 35 years	29	323	1181	10203	0.77	0.51-1.11	0.20
Maternal age above 40 years	***	***	***	***	0.99	0.27-2.14	0.83
Paternal age under 20 years	5	342	46	11186	3.56	1.22	0.0075
Paternal age under 25 years	47	300	1074	10158	1.48	1.07-2.01	0.014
Paternal age under 30 years	153	194	4263	6969	1.30	1.04-1.60	0.021
Paternal age above 35 years	77	270	2732	8500	0.88	0.68-1.14	0.36
Paternal age above 40 years	28	319	958	10274	0.90	0.27-2.14	0.83
II)Association test within group of all individuals with an ADHD diagnosis (including cases of comorbid ASD&ADHD)							
Phenotype	**CNV carriers (n=)**		**Non-CNV carriers (n=)**		**OR**	**CI**	**P-value**
	Expressing phenotype (n=)	**Not expressing phenotype (n=)**	**Expressing phenotype (n=)**	**Not expressing phenotype (n=)**			
ADHD	401	332*	13075	18745*	1.74	1.50-2.02	2.16. e-13
* Comparison of number of CNV carriers in the Danish ASD case group vs. Danish control group							
1)Psychiatric disorders and behavioural psychiatric manifestations							
ASD	83	318	2815	10260	0.95	0.74-1.21	0.69
Schizophrenia or psychotic disorder	28	373	827	12248	1.11	0.73-1.61	0.60
Mood disorder	37	364	1699	11376	0.68	0.48-0.94	0.028
Anxiety	27	374	1381	11694	0.61	0.40-0.89	0.015
OCD	6	395	537	12538	0.36	0.14-0.72	0.012

Intellectual disability (IQ<70)	76	325	1380	11695	1.97	1.52-2.54	1.61e-07
Anorexia nervosa	5	396	277	12798	0.58	0.201-1.21	0.24
	Mean age of diagnosis of all affected individuals (years)		**Mean age of diagnosis of CNV carriers (years)**		**OR****	**CI**	**P-value**
Age of comorbid ASD diagnosis	11.0		10.0		1.08	0.70-1.71	0.74
Age of comorbid ID diagnosis	11.4		12.4		0.62	0.39-0.98	0.041
**OR for being a CNV carrier associated with obtaining the diagnosis earlier than the mean age of diagnosis of all ADHD-diagnosed individuals							
2)Developmental disorders							
Phenotype	**CNV carriers (n=)**		**Non-CNV carriers (n=)**		**OR**	**CI**	**P-value**
	Expressing phenotype (n=)	**Not expressing phenotype (n=)**	**Expressing phenotype (n=)**	**Not expressing phenotype (n=)**			
Developmental coordination disorder	12	389	321	12754	1.22	0.65-2.10	0.50
Developmental delay	30	371	749	12326	1.32	0.89-1.91	0.14
3)Growth and dysmorphic features							
Craniofacial dysmorphism	11	390	233	12842	1.55	0.79-2.73	0.16
Congenital malformation	22	373	514	12429	1.37	1.02-1.79	0.029
Malformation at birth in heart/major	9	392	235	12840	1.23	0.59-2.32	0.51
Other congenital abnormality	62	339	1531	11544	1.38	1.04-1.80	0.022
Microcephaly	10	373	166	5591	0.90	0.44-1.64	0.76
Macrocephaly	5	163	163	5636	0.94	0.33-2.09	0.89
Apgar 5 score under 7	***	***	***	***	1.01	0.40-2.10	0.97
Premature birth (under 37 weeks of gestation)	25	373	866	12116	0.94	0.61-1.38	0.76
Small for gestational age (under 10th percentile)	48	347	1653	11253	1.25	0.91-1.68	0.15
***Number of individuals not presented, when the group is constituted by four or less than four individuals							
4)Somatic comorbidities							
Epilepsy/febrile seizures	48	353	1124	11951	1.44	1.05-1.95	0.019
GI infection	50	351	1512	11563	1.09	0.80-1.46	0.58
Kidney/urinary tract infection	9	392	460	12615	0.63	0.30-1.15	0.17
Unspecified infection	144	257	4345	8730	1.11	0.91-1.38	0.26
Unspecified autoimmune disease	8	393	210	12865	1.24	0.56-2.38	0.54
5)Family history and pattern of inheritance							
Mother or father with ID	***	***	***	***	2.66	0.80-6.57	0.060
Mother or father with epilepsy or febrile seizures	52	346	1401	11507	1.23	0.91-1.65	0.16

Mother with any diagnosis of psychiatric disorder	103	295	3316	9606	1.01	0.80-1.27	0.92
Father with any diagnosis of psychiatric disorder	84	314	2404	10506	1.17	0.91-1.49	0.21
Maternal age under 20 years	19	382	642	12433	0.96	0.58-1.49	0.88
Maternal age under 25 years	139	262	3981	9094	1.21	0.98-1.49	0.072
Maternal age under 30 years	281	120	8841	4234	1.12	0.91-1.40	0.30
Maternal age above 35 years	25	376	880	12195	0.92	0.60-1.36	0.70
Maternal age above 40 years	***	***	***	***	0.64	0.11-2.04	0.54
Paternal age under 20 years	9	389	159	12759	1.86	0.87-3.45	0.074
Paternal age under 25 years	74	324	2137	10781	1.15	0.89-1.48	0.28
Paternal age under 30 years	192	206	6227	6691	1.01	0.82-1.22	0.99
Paternal age above 35 years	68	330	2284	10634	0.96	0.73-1.24	0.76
Paternal age above 40 years	23	375	726	12192	0.89	0.11-2.04	0.54
III) Supplementary association analyses							
Phenotype	**CNV carriers (n=)**		**Non-CNV carriers (n=)**		**OR**	**CI**	**P-value**
	Expressing phenotype (n=)	**Not expressing phenotype (n=)**	**Expressing phenotype (n=)**	**Not expressing phenotype (n=)**			
1)In ASD-cases only (cases of comorbid ADHD&ASD excluded)							
Developmental coordination	14	255	231	8338	1.97	1.08-3.32	0.016
2)In ADHD-cases only (cases of comorbid ADHD&ASD excluded)							
ID (IQ<70)	55	263	906	9354	2.22	1.59-2.89	4.33e-07
Mild ID (50<IQ<70)	41	277	719	9541	1.96	1,34-2,78	8.48e-05
Moderate ID	9	309	180	10080	1.63	0.76-3.03	0.158
Severe ID	NA						
Epilepsy/febrile seizures	36	282	826	9431	1.45	1.00-2.04	0.039
3)Younger parental age in the population control group							
Maternal age under 20 years	15	317	453	18292	1.91	1.08-3.12	0.016
Paternal age under 20 years	5	325	106	18517	2.69	0.94-5.98	0.032

Supplements

Supplementary table 1: Phenotypic features reported in relation to at least two of the twelve selected pathogenic CNVs from the OMIM database across

1)Psychiatric disorders and behavioural psychiatric manifestations											
Reported phenotypic feature	**1q21.12 del.**	**1q21.2 dupl.**	**15q11.2 del.**	**15q11.2 and 15q13 dupl.**	**15q13.2.3 del.**	**16p11.2 del.**	**16p11.2 dupl.**	**17q12 del.**	**17q12 dupl.**	**22q11 del.**	**22q11 dupl.**
ADHD	X	X 29%	X		X 9/14 carriers //6%	X	X			X	
ASD	X P-value: 1.67. 10(-4)	X 41% //50%	X	X	X 6/10 // 11%	X Association between locus and ASD: P-value: 0,039 // 0,002		X		X	
OCD			X		X						
Aggressive behaviour	X				X						
Happy demeanour			X		X 9/14 carriers						
Anxiety	X			X		X	X				
Mood disorder	X 26%			X	X 11%						
ID /cognitive deficits/ decreased IQ	X 16/21 carriers // 76%	X 29% // 62,5%	X	X 75%	X 12/14 carriers // 58%	X	X	X ¾ carriers	X	X	**X**
SZ	X P-value: 1.67. 10(-4)	X Enrichment of this dupl. in cohorts of SZ pts: P-value: 0,015	(MRI showing changes in grey matter corresponding to state of early psychosis)	X	X 11%	X Association between locus and SZ: P-value: 0.039		X		X OR: 1.84, P-value <0,03	
Anorexia nervosa or restrictive eating behaviour				X			X				
Early onset of any diagnosis above			X Case: ID at 3.5 years old	X Case: develop-mental regression at 2.5 years old	X Average age of psychiatric diagnosis 6.6 years			X Signs of ASD under 2 years among 3 cases			

2) Developmental disorders											
	1q21.12 del.	**1q21.2 dupl.**	**15q11.2 del.**	**15q11.2 – 15q13 dupl.**	**15q13.2.3 Del.**	**16p11.2 Del.**	**16p11.2 dupl.**	**17q12 del**	**17q12 dupl.**	**22q11 del**	**22q11 dupl.**
Delay of motor development / poor motor skills	X		X 36%	X	x	X 50%					X
Learning difficulties	X		X 90%	X	X					X	X
Delayed speech			X	X	X 16%	X	X	X		X	X
Developmental delay	X	X	X 59%	X	X 12/14 carriers	X OR, 8.64; P-value: 6.34e(-10))	X	X ¾ cases		X	X
3) Growth and dysmorphic features											
	1q21.12 del.	**1q21.2 dupl.**	**15q11.2 del.**	**15q11.2 – 15q13 Dupl.**	**15q13.2.3 Del.**	**16p11.2 Del.**	**16p11.2 Dupl.**	**17q12 del**	**17q12 dupl.**	**22q11 del**	**22q11 dupl.**
Facial dysmorphic feature	X		X		X	X	X	X	X	X	X
Skeletal abnormalities or malformations of other organ systems	X 17/21 carriers-81%	X 63%	X		X 2,4% (based on 246 cases)	X Rare severe cases// 30% reported from other study	X	X (rare)	X	X	X
Malformation in heart/veins	X 6/21 carriers	X Enrichment of this dupl. in cohorts of tetralogy of fallot P-value=0,004	X 1/9 carriers		X Cardiac defects reported in other study (17%)	X Pt. Died at 5 months, Tetralogy of Fallot			X	X 56% Tetralogy of Fallot	X
Microcephaly	X (or relative microcephaly) 14/21 carriers – 67%				X	X	X			X	X
Macrocephaly		X 50% (or relative macro-cephaly)			X		X	X			
Hypotonia	X	X	X	X	X 47%				X		X 43%

Height	Short stature 50%	Short stature 27%			Short stature 24%		Reduced height	X 4/4 carriers		X	Increased height
Weight				Under-weight	Over-weight / under-weight	Severe over-weight, early onset	Under-weight			Delayed growth	Overweight
Birth- or pregnancy complications				Low birth-weight, 2 cases	Placenta previa, 4 cases					Small for gestational age Low gestational age (27-36 weeks) Preeclam-psia Polyhydramnios	Premature birth
4) Somatic comorbidities											
	1q21.12 del.	**1q21.2 dupl.**	**15q11.2 Del.**	**15q11.2 – 15q13 Dupl.**	**15q13.2.3 Del**	**16p11.2 Del**	**16p11 Dupl.**	**17q12 del**	**17q12 dupl.**	**22q11 del**	**22q11 dupl.**
Epilepsy / seizures / abnormal EEG	X	X	X 2/9 carriers // OR: 4,9 carriers vs controls p-value = 4.2 x 10(-4)	X 15-30% seizures // 20-50% abnormal EEG	X 7/9 // 28%	X 40%	X	X 2/4 cases	X	X (possibly due to hypocalca-emia)	
Disorder related to the musculo-skeletal system (hypermobility/ connective tissue disorder)	X 5/21 carriers	X		X	X 7/9 carriers	X		X	X	X	
Other somatic comorbidities	X	X	X	X	X	X		X	X	X	
Type	Gastric ulcers (33%),	Scoliosis (27%, carpal tunnel syndrome ..) Gastric ulcers	GI disorders, sleep disorders	Tuberous sclerosis		SCID, immune defects Early death due to numerous malforma tions		Diabetes + diseases and infections of kidney, urinary and GI	malforma tions and diseases of kidney urinary system and GI	Infections of ear, diseases of kidney and urinary system, tumors	
5) Family history and pattern of inheritance **Reported across all loci: variable penetrance**											
	1q21.12 del.	**1q21.2 dupl.**	**15q11.2 Del.**	**15q11.2 15q13 Dupl.**	**15q13.2.3 Del**	**16p11.2 Del**	**16p11 Dupl.**	**17q12 del**	**17q12 dupl.**	**22q11 del**	**22q11 dupl.**
Parents with same CNV, expressing overlapping phenotypic features	X		X		X	X mother with ASD/ID	X		X Father ID		X
Parents with same CNV,	X	X	X	X	X				X		X

unaffected phenotype											
Other comments regarding inheritance				22,6% of cases also present with other genomic variants	75% // 86% inherited, study suggests maternal transmission, male carriers more affected suggested by study		Possibly transmitted from mother, boys more likely to express phenotype		Mother is mosaic of the duplication, daughter is not	Family history of psychiatric disorders, no available genetic data	

Supplementary table 2: Matching of variables in iPSYCH corresponding to the reported phenotypic features from the OMIM database

Phenotypic feature reported in OMIM	Matched variable in iPSYCH	Origin
1) Psychiatric disorders and behavioural psychiatric manifestations		
ASD	ICD-10: F8101	PCR*
ADHD	ICD-10: F9201	PCR*
OCD	ICD-10: F4307	PCR*
Anxiety	ICD-10: F4300	PCR*
Mood disorder	ICD-10: F3000	PCR*
Intellectual disability	ICD-10: F7000	PCR* + LPR**
Schizophrenia or psychotic symptoms	ICD-10: F2000	PCR*
Anorexia nervosa or restrictive eating behaviour	ICD-10: F5100	PCR*
Early onset of any diagnosis above	Age of onset of ASD or intellectual disability Age_F8101 Age_F7000	PCR*
* PCR: Psychiatric disorders and behavioural manifestations reported in OMIM were linked to the patients ICD-10 codes or former ICD-8 codes obtained from the Danish Psychiatric Central Research Register before 31. December 2012 with follow-up until 2016 ** LPR: Danish National Patient Register complete until 2012		
2) Developmental disorders		
Delay of motor development / poor motor skills	Developmental coordination disorder ICD-10: F82	LPR**
Learning difficulties	-	-
Delayed speech	-	-
Developmental delay	ICD-10:R62	LPR**
-unavailable data in iPSYCH corresponding to the reported feature		
3) Growth and dysmorphic features		
Facial dysmorphic feature	Craniofacial dysmorphism ICD-10: Q10-Q18	LPR**
Skeletal abnormalities or malformations of other organ systems	Two matched variables: 1)Congenital malformation 2)Other congenital abnormalities ICD-10: Q00-Q07 or Q20-Q89	1)MBR*** until 2012 2)LPR**
Malformation in heart/veins	Malformation at birth in heart/major veins ICD-10: Q2010	LPR**
Microcephaly**	Head circumference at birth	MBR*** 1997-2012
Macrocephaly**	Head circumference at birth	MBR*** 1997-2012
Hypotonia	Apgar 5 score under 7 *As proxi of hypotonia at birth (threshold also applied by other studies)*[32]	MBR*** 1978-2012

Height	-	-
Weight	-	-
Birth- or pregnancy complications	Two matched variables: 1)Premature birth (before 37 weeks of gestation) 2)Small for gestational age: birthweight under 10th percentile	MBR*** until 2012

*** MBR: Medical Birth Registry

****Age of ascertained respective micro- and macrocephaly did not appear from the clinical reports at OMIM. Head circumference at birth obtained from Medical Birth Registry (MBR) for individuals born between 1997 and 2012 was provided in the iPSYCH dataset, hence congenital micro- and macrocephaly was included. Values of 3rd or 97th centile of head circumference according to sex and gestational age was provided by the *International standards for newborn weight, length and head circumference by gestational age and sex: the Newborn Cross-Sectional Study of the INTERGROWTH-21st Project.* [33]. We observed a discrepancy between the 3rd and 97th percentiles measures provided from the *INTERGROWTH-21st Project* and our data in iPSYCH for individuals born at term. Thus, we adjusted the measures of micro- and macrocephaly by subtracting the difference between these two percentiles. It was not possible to assert micro- and macrocephaly for individuals with a gestational age above 42 weeks, since no comparable standard values were provided from the *INTERGROWTH-21st Project*.

4) Somatic comorbidities

Epilepsy / seizures / abnormal EEG	Common variable based on the following ICD-10 codes: G4010 = epilepsy or G4020 = status epilepticus or R5010 = febrile seizures	LPR 1977-2012
Disorder related to the musculo-skeletal system (hypermobility/connective tissue disorder)	-	-
Other somatic comorbidities**	1)Gastrointestinal infection 2)Urogenital infection 3) Unspecified infection (viral or bacterial) 4) Unspecified autoimmune disease	LPR 1977-2012

**** *Reports of infections were linked to available information about somatic hospital contacts from LPR (inpatient and outpatient from 1977 to 2012) and grouped under the variables "Gastrointestinal infections", "Urogenital infections", "Unspecific infections" (including any viral or bacterial infection) and "Nonspecific autoimmune diseases" (the later including i.e. diabetes and thyroiditis, reported for two CNVs).*

5) Family history and pattern of inheritance

Parents expressing overlapping phenotypic features	Three matched variables: 1) Parent with intellectual disability ICD-10: F7000 2) Parent with epilepsy or febrile seizures 3) Parent with any psychiatric diagnosis (ICD-10)	LPR**
Parental age	1) Maternal age at birth 2) Paternal age at birth	MBR***

References

[1] Lai, M. et. al., *Autism*, Lancet, 2014, **383**: p. 896-910.

[2] Polanczyk, G., et. al., *The worldwide prevalence of ADHD: a systematic review and metaregression analysis,* Am J Psychiatry, 2007, **164**(6): p. 942-8.

[3] Loomes, R., et al., *What Is the Male-to-Female Ratio in Autism Spectrum Disorder? A Systematic Review and Meta-Analysis*, J Am Acad Child Adolesc Psychiatry, 2017, **56**(6): p.466-474.

[4] Ramtekkar, U. P., et al., *Sex and age differences in Attention-Deficit/Hyperactivity Disorder symptoms and diagnoses: Implications for DSM-V and ICD-11*, J Am Acad Child Adolesc Psychiatry, 2010, **49**(3): p.217-28, e1-3.

[5] American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders* (*DSM–5*), 2014.

[6] Faraone, S.V., et al., *Genetics of attention deficit hyperactivity disorder*, Molecular Psychiatry, 2019, **24:** p. 562–575.

[7] Tick, B., et al., *Heritability of autism spectrum disorders: a meta-analysis of twin studies,* J Child Psychol Psychiatry, 2016, **57:** p. 585–95.

[8] Anagnostou, E., et al., *Autism spectrum disorder: advances in evidence-based practise*, CMAJ, 2014, **186:** p.509-19.

[9] Kirow, G., et al., *CNVs in neuropsychiatric disorders,* Human Molecular Genetics, 2015, Vol. **24,** No. R1: R45–R49.

[10] Gudmundsson, O., et al., *Attention-deficit hyperactivity disorder shares copy number variant risk with schizophrenia and autism spectrum disorder*, Translational Psychiatry, 2019, **9**(1): Article number: 258.

[11] Torres, M., et. al., *Recurrent copy number variations as risk factors for neurodevelopmental disorders: critical overview and analysis of clinical implications*, J Med Genet, 2016, **53**(2): p. 73-90.

[12] Gamazon, E.R., et. al., *The impact of human copy number variation on gene expression*, Brief Funct Genomics, 2015, **14**(5): p. 352–357.

[13] Shprintzen, R.J., et. al., *A new syndrome involving cleft palate, cardiac anomalies, typical facies and learning disabilites: velo-cardio-facial syndrome,* Cleft Palate J., 1978, **15**: p. 56–62.

[14] Driscoll, D.A., et. al., *Deletions and microdele- tions of 22q11.2 in velo-cardio-facial syndrome*, 1992, Am. J. Med. Genet., **44**, p: 261–268.

[15] Vinas-Jornet, M., et al., *High Incidence of Copy Number Variants in Adults with Intellectual Disability and Co-morbid Psychiatric Disorders,* Behav Genet, 2018, **48**(4): p.323-336.

[16] Qiao, Y., et al., *Phenomic determinants of genomic variation in autism spectrum disorders*, J Med Genet, 2009, **46**(10): p.680-8.

[17] Sullivan, P.F., et. al., *Genetic architectures of psychiatric disorders: the emerging picture and its implications,* Nat Rev Genet. 2012, **13**(8): p. 537-51.

[18] Martin, C., et. al., *Identification of Neuropsychiatric Copy Number Variants in a Health Care System Population,* JAMA Psychiatry, 2020, **77**(12) p: 1–10**.**

[19] Kearney, H.M., et. al., *American College of Medical Genetics standards and guidelines for interpretation and reporting of postnatal constitutional copy number variants, Genet Med,* 2011, **13**(7):p.680-5.

[20] Pedersen, C.B., et al., *The iPSYCH2012 case-cohort sample: new directions for unravelling the genetic and environmental architecture of severe mental disorders*, Mol Psychiatry, 2018. **23**(1): p. 6-14.

[21] Bybjerg-Grauholm, J., et. al., *The iPSYCH2015 Case-Cohort sample:updated directions for unravelling genetic and environmental architectures of severe mental disorders,* In Review

[22] Olsen, L., et. Al., *Rearrangements in the 22q11.2 Region: Prevalence and Population-Based Risk for Neuropsychiatric and Developmental Disorders,* Lancet Psychiatry, 2018, **5**(7): p 573-580.

[23] *The Online Mendelian Inheritance in Man database*, McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, under the direction of Dr. Ada Hamosh

[24] Harrell, F., *Regression coefficients and scoring rules,* J Clin Epi- demiol, 1996, **49**: p. 819.

[25] Naylor., D., et al., *Letter to editors: Clinical prediction rules*, J Clin Epidemiol, 1997, **Vol. 5.** no. 6: p. 743-744.

[26] Sanchez, X.C., et. al., *Prevalence, risk of disease, fertility and mortality of copy number variants in the Danish population: A case-cohort study,* In Review

[27] Green, E. K., et al., *Copy number variation in bipolar affective disorder,* Mol. Psychiatry, 2015, **21:** p89-93.

[28] McGrath, L.M., et. al., *Copy Number Variation in Obsessive-Compulsive Disorder and Tourette Syndrome: A Cross-Disorder Study,* J Am Acad Child Adoles Psychiatry, 2014, **53**(8): p.910-919.

[29] Kendall, K.M, et. al., *Association of Rare Copy Number Variants With Risk of Depression, JAMA Psychiatry*, 2019, **76**(8): p.818–825.

[30] Mosca, S.J., et. al., *Copy-number variations are enriched for neurodevelopmental genes in children with developmental coordination disorder*, J Med Genet, 2016, **53**(12): p. 812-19.

[31] Malaspina D., et. al., *Paternal factors and schizophrenia risk: de novo mutations and imprinting,* Schizophr Bull, 2001, **27**: p.379–93.
[32] Hultman., C., et al., *Prenatal and perinatal risk factors for schizophrenia, affective psychosis, and reactive psychosis of early onset: case-control study*, BMJ, 1999, **318**(7181): p.421-426.

[33] Villar., J., et al., *International standards for newborn weight, length, and head circumference by gestational age and sex: the Newborn Cross-Sectional Study of the INTERGROWTH-21st Project*, Lancet, 2014, **384**(9946): p.857-68.