Using common genetic variation to examine phenotypic expression and risk prediction in 22q11.2 Deletion Syndrome

Robert W. Davies[1,2]*, Ph.D, Ania M. Fiksinski[3,4]*, M.Sc., Elemi J. Breetvelt[5], M.D., Ph.D., Nigel M. Williams[6], Ph.D., Stephen R. Hooper[7], Ph.D., Thomas Monfeuga[6], Ph.D., Anne S. Bassett[4,8,9], M.D., FRCPC, Michael J. Owen[6], M.D., Ph.D., Raquel E. Gur[10], M.D., Ph.D., Bernice E. Morrow[11], Ph.D., Donna M. McDonald-McGinn[12,13], M.Sc., LCGC., Ann Swillen[14,15], Ph.D., Eva W. C. Chow[4,8], M.D., Marianne van den Bree[6], Ph.D., Beverly S. Emanuel[12], Ph.D., Joris R Vermeesch[14], Ph.D., Therese van Amelsvoort[16], M.D., Ph.D., Celso Arango[17], M.D., Ph.D., Marco Armando[18], M.D., Ph.D., Linda E. Campbell[19], Ph.D., Joseph F. Cubells[20,21], M.D., Ph.D., Stephan Eliez[18], M.D., Sixto Garcia-Minaur[22], M.D., Doron Gothelf[23,24], M.D., Wendy R. Kates[25], Ph.D., Kieran C. Murphy[26], M.D., Ph.D., Clodagh M. Murphy[27], MD, Ph.D., Declan G. Murphy[27], M.D FRCPsych, Nicole Philip[28,29], M.D., Gabriela M. Repetto[30], M.D., Vandana Shashi[31], MBBS., M.D., Tony J Simon[32], Ph.D., Damiàn H. Suñer[33], Ph.D., Stefano Vicari[34], M.D., Stephen W. Scherer[36], Ph.D., International 22q11.2 Brain and Behavior Consortium[#], Carrie E. Bearden[35], Ph.D., Jacob A.S. Vorstman[3,36,37], M.D., Ph.D.

* These authors contributed equally to this work
# Group authorship; authors to be listed in the acknowledgement section at the end of this paper

Affiliations:
1. Program in Genetics and Genome Biology and The Centre for Applied Genomics, The Hospital for Sick Children, Toronto, Ontario, Canada.
2. Department of Statistics, University of Oxford, 24-29 Giles' St., Oxford, OX1 3LB, United Kingdom
3. Department of Psychiatry, Brain Center, University Medical Center Utrecht, the Netherlands.
4. The Dalglish Family 22q Clinic, Clinical Genetics Research Program, Centre for Addiction and Mental Health, Toronto, Canada.
5. Department of Psychiatry, The Hospital for Sick Children, Toronto, Ontario, Canada
6. MRC Centre for Neuropsychiatric Genetics and Genomics, Division of Psychological Medicine and Clinical Neurosciences, School of Medicine, Cardiff University, Cardiff, United Kingdom.
7. Department of Allied Health Sciences, School of Medicine, University of North Carolina-Chapel Hill, North Carolina, United States of America.
8. Department of Psychiatry, University of Toronto, Ontario, Canada.
9. Toronto General Hospital, University Health Network, Toronto, Ontario, Canada.
10. Department of Psychiatry and Lifespan Brain Institute, Penn Medicine - CHOP, University of Pennsylvania, Pennsylvania, United States of America.
11. Department of Genetics, Albert Einstein College of Medicine, Bronx, New York, United States of America.

12. Division of Human Genetics and 22q and You Center, the Children's Hospital of Philadelphia, Pennsylvania, United States of America.
13. Department of Pediatrics, Perelman School of Medicine of the University of Pennsylvania, Philadelphia, United States of America.
14. Center for Human Genetics, University of Leuven (KU Leuven), Belgium.
15. Department of Human Genetics KU Leuven, Belgium.
16. School Mental Health & Neuroscience Maastricht University, the Netherlands.
17. Department of Child and Adolescent Psychiatry, Hospital General Universitario Gregorio Marañón, IiSGM, CIBERSAM, School of Medicine, Universidad Complutense, Madrid, Spain
18. Developmental Imaging and Psychopathology, Department of Psychiatry, University of Geneva, Switzerland.
19. School of Psychology, University of Newcastle, Newcastle, Australia.
20. Department of Human Genetics, Emory University School of Medicine, Atlanta, Georgia, United States of America.
21. Emory Autism Center, Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, Atlanta, Georgia, United States of America.
22. Institute of Medical and Molecular Genetics (INGEMM), La Paz University Hospital, Madrid, Spain.
23. The Child Psychiatry Division, Edmond and Lily Safra Children's Hospital, Sheba Medical Center, Tel Hashomer, Israel.
24. Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel.
25. Department of Psychiatry and Behavioral Sciences, SUNY Upstate Medical University, New York, United States of America.
26. Department of Psychiatry, Royal College of Surgeons in Ireland, Beaumont Hospital, Dublin, Ireland
27. Department of Forensic and Neurodevelopmental Sciences, Institute of Psychiatry, Psychology & Neuroscience (IoPPN), King's College London, United Kingdom.
28. Département de Génétique Médicale, APHM, CHU Timone Enfants, Marseille, France.
29. Aix Marseille Université, MMG, INSERM, Marseille, France.
30. Centro de Genética y Genómica, Facultad de Medicina, Clínica Alemana Universidad del Desarrollo, Santiago, Chile.
31. Department of Pediatrics, Division of Medical Genetics, Duke University School of Medicine, North Carolina, United States of America.
32. MIND Institute and Department of Psychiatry and Behavioral Sciences, University of California Davis, Sacramento, California, United States of America.
33. Laboratorio Unidad de Diagnóstico Molecular y Genética Clínica, Hospital Universitari Son Espases, Palma de Mallorca, Spain.
34. Child and Adolescence Neuropsychiatry Unit, Department of Neuroscience, IRCSS Bambino Gesù Children's Hospital of Rome, Italy.
35. Departments of Psychiatry and Biobehavioral Sciences and Psychology, Semel Institute for Neuroscience and Human Behavior. University of California, Los Angeles, California, United States of America.

36. Program in Genetics and Genome Biology, SickKids Research Institute, Toronto, Ontario, Canada.
37. Department of Psychiatry, The Hospital for Sick Children, University of Toronto, Ontario, Canada.

Corresponding Author
Jacob A.S. Vorstman, MD, PhD
Program in Genetics and Genome Biology, SickKids Research Institute,
686 Bay str, Office 12.9702,
M5G 0A4, Toronto, Ontario
Canada
Phone +1 289 305 5587
        +1 437 776 4371
Email: jacob.vorstman@sickkids.ca

Word count:

Figures: 2

Tables: 2

Introductory Paragraph (172)

The 22q11.2 deletion syndrome (22q11DS) is associated with a 20 – 25% risk for schizophrenia.[1,2] We examined phenotypic and genetic data in a cohort of 962 individuals with 22q11DS to evaluate the shared genetic basis between schizophrenia and schizophrenia-related early trajectory phenotypes: subthreshold symptoms of psychosis, low baseline intellectual functioning, and cognitive decline. We examined the association of these phenotypes with two polygenic scores, derived for schizophrenia and intelligence, and evaluated their use for individual risk prediction in 22q11DS. These polygenic scores were not only associated with schizophrenia and baseline IQ, respectively, but schizophrenia polygenic score was also significantly associated with cognitive (verbal IQ) decline and nominally associated with subthreshold psychosis. Further, comparing the tail-end deciles of the schizophrenia and IQ polygenic score distributions, 33% versus 9% of 22q11DS subjects had schizophrenia, and 63% versus 24% had intellectual disability. Collectively, these data show both a shared genetic basis for schizophrenia and schizophrenia-related phenotypes, and highlight the future potential of polygenic scores for risk stratification among individuals with highly, but incompletely, penetrant genetic variants.

Main Text Word count (intro + result + discussion) 2,557

While schizophrenia (SZ) is typically diagnosed in late adolescence or early adulthood, it is now well established that the first psychotic episode is in fact a manifestation of an advanced stage of this illness.[3] Early behavioral, cognitive and neuroanatomic changes are measurable prior to the first psychotic episode.[4-8] Both lower cognitive ability early in life (the estimated premorbid deficit is 8 IQ points[9]), as well as cognitive decline in early adolescence (estimated IQ-change equal to -1.09 standard deviation[10]), are associated with schizophrenia, with effect sizes in the range of 0.4 to 0.5.[9-15] In addition, subthreshold psychotic symptoms in youth also index increased risk for schizophrenia.[16-18] These observations raise an important question: Do early cognitive phenotypes and subthreshold symptoms of psychosis share a substantial genetic basis with either schizophrenia or intellectual ability?

Early schizophrenia-related phenotypes and trajectories are difficult to study, requiring longitudinal follow-up of large cohorts to capture a sufficient number of schizophrenia cases. Populations at increased risk for schizophrenia, such as patients with 22q11DS, facilitate such studies.[19] Findings from 22q11DS studies reproduce observations related to schizophrenia in the general population, thereby supporting 22q11DS as a genetic model of schizophrenia, including its early trajectory.[20] In 22q11DS, as in the general population, subthreshold psychotic symptoms, low baseline intellectual ability and increasing cognitive deficits over time, particularly in verbal IQ, are all associated with increased risk of subsequent psychotic illness. [21]

A large fraction of the heritability of schizophrenia comes from a polygenic burden of multiple common variants, each of small effect.[22,23] Increasingly, polygenic scores derived from genome-wide association studies (GWAS) have been used to study the genetic relationship between phenotypes.[24] For schizophrenia, a polygenic score using recent GWAS explains up to 7% of the variance on the liability scale.[23] Similarly, polygenic scores for general cognitive function, or proxies thereof, explain 2.5 - 4.3% of its variance.[25,26] Polygenic scores can also be used for phenotype prediction.[27] In the general population, they are not yet particularly effective as individual risk predictors[28] given the relatively low population prevalence of phenotypes such as schizophrenia and intellectual disability (ID), and the still modest effect sizes conferred by polygenic scores.[29] However, in high-risk populations such as 22q11DS, the same effect size acts upon a higher baseline prevalence (e.g., 25% for schizophrenia), which may allow for more substantial differences in absolute risk.[30]

The International 22q11.2 Deletion Syndrome Brain Behavior Consortium (IBBC) has assembled the largest genotype-phenotype dataset of individuals with 22q11DS.[20] Previously, the IBBC has reported on genetic associations of both common and rare variants in 520 individuals with 22q11DS, exclusively focusing on schizophrenia[31]. The current study presents several novel analyses, conducted in a substantially larger cohort of individuals with 22q11DS (N = 962) and including longitudinal IQ data. Our main objectives were twofold. First, to study the genetic relationship between schizophrenia

and schizophrenia-related phenotypes (i.e., low baseline intellectual ability, cognitive decline and subthreshold positive psychotic symptoms). Second, to examine the use of polygenic scores for schizophrenia and IQ for individual risk prediction of schizophrenia and ID in individuals with 22q11DS.

After applying phenotypic classification and performing genotype quality control, data from 962 IBBC cohort members were available for analysis (**Table 1, Methods**). Within this cohort we distinguished those with Schizophrenia Spectrum Disorder (SSD; N = 207 [20]), subthreshold psychotic symptoms; N = 158), and those with neither phenotype, grouped into "putative controls" (age <25 ("putative" given the typical age at onset of schizophrenia [32]), N = 382) and "definite controls" (age >= 25, N = 215) (**Table 1, Methods**). Subsequently, we refer to all controls regardless age as "merged controls" (N = 597). Baseline Full Scale IQ (FSIQ) was transformed to z-score as previously described,[21] with an average near 0 (0.03; **Table 1**). VIQ decline, operationalized as exceeding –1 SD (binary), occurred in 5.9% of the cohort. We constructed polygenic scores for schizophrenia[23] (PS_SZ) and intellectual ability[33] (PS_IQ) using standard methods (**Methods**), and performed statistical analyses using either linear or logistic regression as appropriate, adjusting for age, sex and the first five principal components from the imputed genotypes (**Methods**). [31]

	SSD	Sub-threshold psychosis	Putative control	Definite control	All	p^c
N max	207	158	382	215	962	N/A
Sex %M	49	49	54	39	49	0.008
Age at last assessment*	31.6 (12.7) [205], {7,64}	17.9 (5) [158] {8,36}	15.2 (4.6) [382] {5,24}	36.8 (9.9) [215], {25,67}	24 (12.4) [960], {5,67}	2.0×10^{-167}
Baseline FSIQ*a	-0.34 (0.87) [145], {-2,2}	0.13 (0.96) [127] {-2,3}	0.07 (1.03) [308] {-3,2}	0.24 (0.85) [120], {-2,2}	0.03 (0.97) [700], {-3,3}	1.8×10^{-6}
Binary VIQ Decline**b	11.9% (59)	5.7% [87]	4.5% [198]	4.7% (43)	5.9% (387)	0.21
Co-morbid mood disorders**	41% (144)	29.2% [154]	16% [363]	38.6% (153)	27.1% (814)	1.7×10^{-10}

Table 1. Clinical characteristics of sample cohort. *quantitative phenotypes are given as mean (SD) [N] {range}, ** binary phenotypes are given as percent true [N], a Baseline FSIQ is given as a z-score using previously defined normalization procedure,[21] b Binary VIQ decline is operationalized as VIQ decline > -1SD (i.e. 1 z-score), as defined by the reliable change index,[34] c p-value is from an ANOVA of phenotype by group

First, we examined known associations between phenotypes and polygenic scores. We observed a significant association between SSD cases versus controls and PS_SZ (N = 802, p = 4.37×10^{-8}, marginal Nagelkerke pseudo-r^2 = 0.053; p-values reported in the text are nominal), and a similar result when including definite controls only (N = 420, p = 1.89×10^{-6}, r^2 = 0.071) (**Table 2, Supplementary Figure 1**), corroborating previous reports from the comprehensive IBBC genetic analyses related to schizophrenia in 22q11DS[31]. We also observed a significant association between baseline FSIQ and PS_IQ (p = 1.08×10^{-7}), and consistent with the known genetic correlation between schizophrenia

and IQ (r2g = -0.234), observed a nominal association between baseline FSIQ and PS_SZ

(p = 0.018) as well as a significant association between SSD and PS_IQ (p = 7.15x10^{-4}).

Then, we assessed relationships between schizophrenia-related phenotypes and the

polygenic scores. We observed a decreasing trend of PS_SZ for phenotypes of SSD

(mean = 0.23), subthreshold psychosis (mean = 0.16), putative controls (mean = -0.05),

and definite controls (mean = -0.27) (**Figure 1**). PS_SZ was nominally significantly higher

in those with subthreshold psychosis compared to the merged control groups (N = 755,

p = 0.0247, r^2 = 0.01, **Table 2, Figure 1**). Numerous post-hoc analyses ruled out some

potential sources of confounding (**Supplementary Note 1**). We also modeled the impact

of future conversion to psychosis in current putative controls and found a moderate

negative correlation between age and PS_SZ, (r2 = -0.2, p =0.048), in line with

expectations that increasing age would be associated with lower likelihood of

developing psychosis (**Supplementary Note 2**). Finally, we observed a significant

association between VIQ decline and PS_SZ (p = 5.09x10^{-3}). Neither the association

between subthreshold psychosis and PS_IQ (p = 0.056), nor between VIQ decline and

PS_IQ (p = 0.658), reached statistical significance.

Dependent variable	IV	N	beta	r^2	p
SSD		802	0.56	0.053	4.37×10^{-8} *
Subthreshold psychosis	PS_SZ	755	0.24	0.01	0.0247
Baseline FSIQ		720	-0.096	0.0077	0.018
VIQ decline		396	0.66	0.051	0.00509 *
SSD		802	-0.30	0.020	7.15×10^{-4} *
Subthreshold psychosis	PS_IQ	755	-0.18	0.0072	0.056
Baseline FSIQ		720	0.20	0.038	1.08×10^{-7} *
VIQ decline		396	-0.096	0.0013	0.658

Table 2. Key regression results. Results are adjusted for standard covariates as described in **Methods**. Beta is the standard regression effect size estimate. r^2 denotes difference between model fit with or without independent variable, using either standard r^2 from linear regression or Nagelkerke pseudo-r^2. Nominal p-values are reported. Asterisk (*) indicates significant result after Bonferroni correction for 8 independent main analyses in this study (two polygenic scores, four phenotypes). IV = Independent Variable.
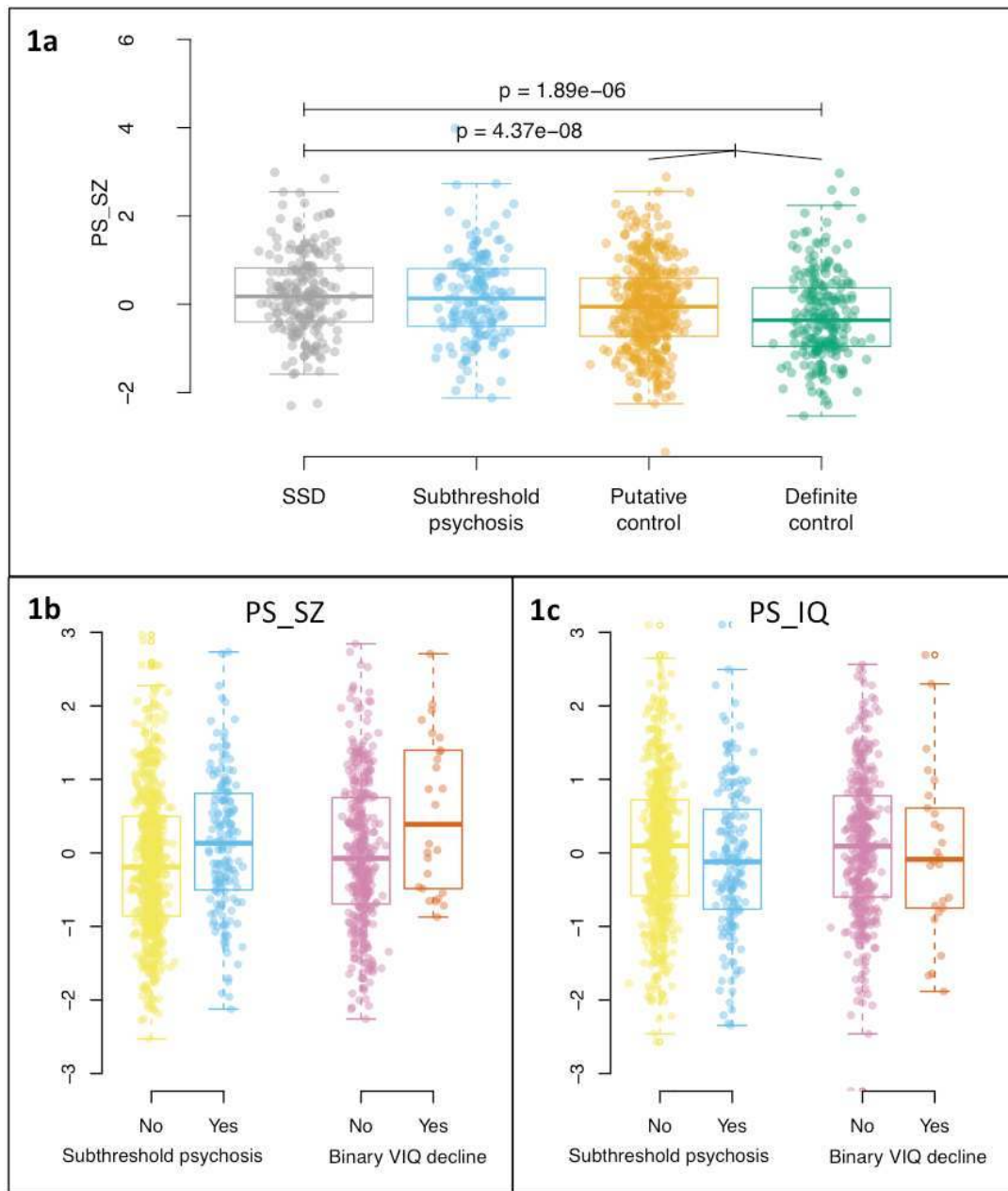
Figure 1. Polygenic scores of schizophrenia (PS_SZ) and intellectual ability (PS_IQ). Per-group individual as well as box-plots highlighting median, interquartile range and 95th percentiles are shown. **a)** Associations of PS_SZ in 22q11DS with SSD, subthreshold psychosis, putative controls and definite controls. Results for PS_SZ and SSD have been reported previously[31], and are included in this figure for completeness. Total sample sizes for the highlighted associations are N=423 (SSD versus putative controls) and N=802 (SSD versus merged controls). **b, c)** PS_SZ (b) and PS_IQ (c) on subthreshold psychosis and verbal IQ decline.

Addressing the second objective of our study, we investigated the extent to which polygenic scores could be used for individualized risk prediction among subjects with 22q11DS. We divided the cohort into quantiles based on polygenic scores, and calculated positive predictive values (PPVs) in each. For SSD, 32% of individuals with scores above the median PS_SZ had SSD (i.e., a PPV of 32%), versus 20% of those with scores below the median (OR = 1.91, 95% CI = [1.38, 2.64], p = 8.4×10^{-5}) (**Figure 2**, **Supplementary Table 1**). Values at the tails showed more extreme differences, with those exceeding the 90th percentile at substantially higher risk (33%) than those in the lowest decile (9.1%). Simulating an effect for the general population, using the observed effect sizes and assuming a general population prevalence of SSD of 1%, generated substantially smaller absolute differences (**Figure 2**).

Similarly, using intellectual ability as a binary outcome (intellectual disability (ID) as IQ < 70), we observed a higher rate of ID among those with a PS_IQ below the median versus above (PPV = 49% vs 34%, OR = 1.85, 95% CI = [1.37, 2.51], p = 7.1×10^{-5}) (**Figure 2, Supplementary Table 2**). This effect is accentuated at the tails, with PPVs of 63% for those in the lowest decile for PS_IQ (i.e. associated with lower IQ in the general population), versus 24% for those in the highest decile of PS_IQ (**Figure 2, Supplementary Table 2**).
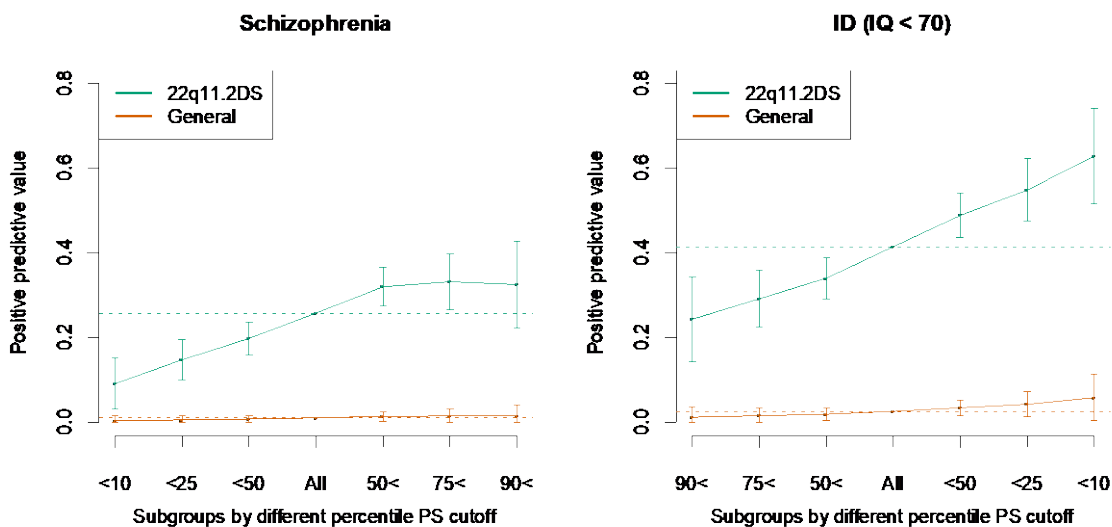
Figure 2. Individual risk prediction. PPVs (y - axis) for SSD (left panel) and ID (right panel) based on various cut-offs of PS_SZ (left panel) or PS_IQ (right panel). Colors differentiate values from the 22q11DS cohort (turquoise) versus values estimated from the general population (orange) given observed prevalences in the population (SSD = 0.01, ID = 0.025; dotted lines) and observed odds ratios. Whiskers represent confidence intervals (+/- 1.96 * standard error).

In this work, we used polygenic scores from large GWAS for schizophrenia and IQ both to better understand the association between schizophrenia and schizophrenia-associated phenotypes, as well as to assess their potential for individual risk prediction. In the first part of the study, we confirmed several results known to occur in the general population, and showed for the first time that known relationships between schizophrenia and IQ extend to individuals with 22q11DS. We observed that a polygenic score for IQ explained ~3.8% of the variance in IQ in 22q11DS, suggesting the previously observed association between parental educational attainment and cognitive outcome in offspring with 22q11DS[35] may be at least partly explained by common variants.

13

In addition, we identified two novel associations between schizophrenia-related phenotypes and schizophrenia. First, we observed a novel association between subthreshold psychosis and PS_SZ. However, this only reached nominal statistical significance, and while post-hoc investigations ruled out several potential confounding sources, further studies are required to provide more certainty in this regard. In addition, we observed a novel, significant association between VIQ decline and PS_SZ, but not PS_IQ, suggesting that common risk variants for schizophrenia contribute to cognitive decline, while common variants associated with cognitive ability might not. A possible implication of these results is that cognitive decline prior to the first psychotic episode may not merely be a risk factor for schizophrenia, as reported previously for 22q11DS[21] and idiopathic schizophrenia,[36,37] but also shares its genetic underpinnings. A previous study in a subset of this cohort showed that cognitive decline preceded the onset of the first psychotic episode by several years,[21] making reverse causation – i.e., cognitive decline as a consequence of psychosis – a less likely explanation. The observed cognitive decline in 22q11DS could be caused by the inability of patients to keep up with peers, or alternatively, represent an absolute loss of cognitive abilities, or a combination of both. The current analyses do not distinguish between these, but prior studies in 22q11DS have found evidence in support of both mechanisms[38,39]. We cannot fully exclude the possibility that the observed cognitive decline could be impacted by the negative effect of psychosis on cognitive testing. However, this is an unlikely explanation given that all study sites refrained from assessing subjects when acutely psychotic, as is common clinical policy. Furthermore, in our data the mean age at IQ assessment is

14

below the age at psychosis onset for both baseline (14.8 and 20.6 years respectively) and longitudinal IQ data (18.2 and 20.3 years respectively). In addition, in those without psychosis, 55% show an IQ decline, versus 45% stable or increase (p = 0.02), indicating that on average, a modest cognitive decline can be observed in 22q11DS regardless of the occurrence of a psychotic disorder, as previously reported.[21] Taken together, our findings are consistent with the notion that disruption of normal cognitive development is a core component of schizophrenia, and investigation of high-penetrance variants for both phenotypes offers important insights into its mechanism.[13]

In the second part of the study, we examined to what extent polygenic scores could be used for individual risk prediction of SSD and ID among individuals with 22q11DS. Whereas in research the existence of association between test and outcome is most relevant, in clinic the positive predictive value (PPV) is key, as it enables stratification of individuals into groups with different outcome probabilities that can inform clinical decision-making. [29,40] Previous studies have shown that high-risk copy number variant (CNV) carriers as a group have increased polygenic scores[41,42], including specifically 22q11DS[31,41], but have not looked at stratification within those groups. Importantly, PPV depends not only on the strength of association, but also on the baseline prevalence. Here we examined risk stratification among individuals with 22q11DS, taking advantage of the higher baseline prevalence of schizophrenia and ID compared to the general population (in our sample, 23% and 41%, respectively). Among those in the highest PS_SZ risk decile, 33% had schizophrenia, versus 9% in the lowest decile. Applying the

same effect sizes to the general population would yield estimates of 1.5% and 0.3%, respectively. Similarly, 63% of those in the lowest PS_IQ decile had ID, versus 24% in the highest decile.

The observed differences between PPVs in our study are similar to those previously reported for BRCA1 and BRCA2 among females for breast cancer risk[30] and males for prostate cancer risk.[43] The concept of using polygenic background to inform individual risk prediction and clinical decision-making is an area of active investigation[44], and is being incorporated into clinical trials for common medical conditions (e.g.,[45]). While our findings highlight the potential clinical utility of polygenic scores in the context of a high-penetrance variant like 22q11DS, the PPVs reported here are not yet sufficient to impact clinical decision-making at present. In addition, while risk prediction enables stratification within high-risk populations, it is important to note that the reduction in risk of those in the lowest risk strata within the 22q11DS population does not bring them to population risk levels. At present, compared to the general population, increased risk for certain outcomes remains a clinical reality for all patients with 22q11DS, regardless of PS results. However, as ever-increasing GWAS size improves the strength of PS associations, we suggest that PS may have clinical utility in risk models in the near future[28], particularly in sub-populations selected for *a priori* increased baseline risk, such as patients with a high-impact mutation like 22q11DS or those with behaviorally defined subthreshold symptoms[46]. Pending more substantial PS effect sizes, as well as robust replication of findings reported here, there are several areas of

potential future clinical utility. For example, in the 22q11DS population, elevated PS_SZ

could be a reason to further intensify monitoring during adolescence, and PS_IQ may

play a role in seeking to prevent misalignment between academic potential and

demands.[47] Taken together, our findings highlight the potential clinical utility of

polygenic scores in the context of a high-penetrance variant.

Further, estimating risk raises important ethical questions, which require careful

consideration. For instance, in the absence of preventative interventions that can alter

outcomes such as schizophrenia or ID, it will be essential to examine the balance

between benefit and potential harm of exposing caregivers and patients to such

information. Studies are required to examine to what extent early risk knowledge can be

used to improve outcomes.[3] Findings like those reported here should prompt a broad

societal discussion about the ethical framework in which they can be used.

While this 22q11DS cohort is the largest ever reported, there are limitations to the work

shown here. Recruitment into the IBBC cohort is not random so there will be

ascertainment biases, which will affect prevalence estimates, but this is not expected to

substantially impact the interpretation of the genotype-phenotype results reported in

the current study (**Supplementary Note 3**). For all analyses, given the current lack of

transferability of polygenic score results across genetic backgrounds[48], and that GWAS

for schizophrenia and IQ are only sufficiently large within European populations to be

powerful, our results were limited to 22q11DS subjects of European descent. Future

large GWAS from diverse backgrounds, and methodological improvements, will allow for analyses in more diverse cohorts. In addition, other uncaptured environmental variables are likely to modulate risk among 22q11DS carriers, and should therefore be included in future studies. Finally, from a multiple testing standpoint, we intentionally restricted the main investigation of schizophrenia and associated phenotypes and polygenic scores to eight. Nonetheless, two of the associations, including one of the novel associations, were only nominally significant, necessitating further investigations for more definitive evidence.

In conclusion, common variants associated with schizophrenia risk and IQ variability in the general population modify expression of these phenotypes in 22q11DS. Verbal IQ decline, and subthreshold psychosis at least partly share genetic underpinnings with schizophrenia, highlighting shared causal pathways between these phenotypes. Furthermore, in 22q11DS carriers polygenic scores enable stratification into high and low risk groups substantially in excess of what would be found in a general population setting. We suggest that in populations with high-risk rare pathogenic genetic variants such as 22q11DS, this approach is nearing a level of differentiation required for clinical utility.

METHODS

Dataset

All individuals in this study were carriers of the 22q11.2 deletion, confirmed by

Multiplex Ligation-dependent Probe Amplification[49] as described previously[31]. All

participants were recruited by one of 22 international IBBC sites (total N = 1,789). Local

research ethics boards provided appropriate study approval at all sites, and all

individuals provided written informed consent regarding participation in this research.

Psychiatric assessment

Psychiatric assessment was performed using standardized semi-structured interviews,[20]

leading to a categorization of each participant in one of the following subgroups:

schizophrenia spectrum disorder (SSD), subthreshold psychosis, putative control, and

definite control. SSD included schizophrenia, schizoaffective disorder and related

psychotic disorders such as delusional disorder or psychotic disorder not otherwise

specified, all in accordance with DSM-IV criteria, based on data obtained by semi-

structured in-person interviews at each site (see[20] for case consensus procedures). Any

individual who had never met criteria for any psychotic disorder diagnosis, but had

endorsed clinically significant positive symptoms at any timepoint, was included in the

subthreshold psychosis group. Supporting scores from various standardized assessment

methods used across sites included symptom scores in the moderate to severe range,

i.e., scores of 3-5 on the Structured Interview for Prodromal Syndromes, (SIPS[50]), or

above 2 on the Comprehensive Assessment of At-Risk Mental States (CAARMS[51]), or of 2

or higher (probable or definite) on any of the positive symptoms on the Schedule for Affective Disorders and Schizophrenia for School-Age Children (K-SADS[52]). 22q11DS individuals without a lifetime diagnosis of any psychotic disorder and who had never endorsed subthreshold positive psychotic symptoms were considered controls. Given that the risk of developing schizophrenia is most elevated until age 25,[53] those younger than 25 years at the most recent assessment were considered "putative controls", while those aged 25 years or older at most recent assessment were classified as "definite controls" (**Supplementary Figure 2,** and demographics in **Supplementary Table 3**).

IQ values and definition of cognitive decline

We previously found baseline Full Scale IQ (FSIQ) to be a significant risk factor for subsequent schizophrenia spectrum disorder in 22q11DS, while the strongest effect size for cognitive decline was observed for verbal IQ (VIQ) [21]. To remain consistent with our prior observations, we considered first available FSIQ as a measure of baseline intellectual ability and change in VIQ between the first and last available measurement as an index of cognitive decline. Given the moderate cognitive decline that occurs, on average, in this population,[38] we calculated standardized values (z-scores) derived from the normative chart on which the average IQ trajectory for the 22q11DS population is mapped. Thus, a decline represents a negative deviation from the expected decline in this population.

In seeking to operationalize a cognitive decline as a binary variable, we sought a cut-off between lenient (i.e. requiring less severe decline, but could introduce too much noise), and conservative (more severe decline, but could reduce *a priori* power). This task is further complicated by potential error variance inherent in the data collection across multiple sites, different versions of the Wechsler Intelligence Scale and different age groups.

We initially performed our analysis using a threshold of more than -0.5 SD as the cut-off for verbal IQ decline. Using this cut-off, the observed association with PS_SZ was not statistically significant (N = 396, p = 0.22, r^2 = 0.006). Based on the literature on the reliable change index,[34] we subsequently revised our definition of significant change to a more stringent threshold. To minimize the chance that any observed decline was due to chance, we conservatively used the lower boundary of the reliable change index as the cut-off, i.e. defining Verbal IQ decline as a binary variable operationalized as any negative change in z-scores exceeding 1 SD difference.

Genotyping methods and principal components analysis

For a total of 1,789 individuals with a 22q11.2 deletion, phenotypic data were collected in a central consortium database and available DNA samples were genotyped at Albert Einstein College of Medicine, in New York, using Affymetrix Human 6.0 microarrays. We generated imputed genotypes from genotyping microarray data using standard methodological approaches as described elsewhere[31]. After imputation, genotype data for 992 individuals and 6,354,586 autosomal single nucleotide polymorphisms (SNPs)

21

were available for inclusion. We retained 4.0 million (M) SNPs, which had minor allele frequency (MAF) > 10% and were not in the major histocompatibility complex (MHC; chromosome 6, 26-34 Mbp) or in the 22q11.2 region (chromosome 22, 18,820,303 to 21,489,474 bp).

For principal component analysis (PCA), we then intersected this with the available GWAS SNPs described below to yield 3.2M SNPs. We ran PCA on the 992 individuals at the 3.2M SNPs using PLINK version 1.9 release 180612,[54] which revealed between-cohort differences matching geographic ascertainment locations, but no obvious outliers for quality control (QC) or non-European ancestry (**Supplementary Figure 3**). Of the 992 individuals who met criteria for subsequent analysis, 27 did not fall into one of the four pre-specified phenotype groups: 21 individuals who were diagnosed with a mood disorder with psychotic features, but who did not meet criteria for any non-affective psychotic disorder, and 6 individuals with insufficient phenotypic data. We further removed three samples that overlapped with the CLOZUK cohort[23], which was a component of the PGC schizophrenia GWAS, yielding a total sample for analysis of n=962 (**Supplementary Figure 4**).

Polygenic score construction

We sought out large GWAS that would enable us to generate maximally predictive polygene scores for schizophrenia and IQ. For SSD, we used published summary statistics from a schizophrenia GWAS from the Psychiatric Genomics Consortium (PGC)

(max N samples = 77,096).[23] For intellectual ability / IQ, we used results from Davies *et al.*[25] from a GWAS for a general intelligence factor, or "g-factor".[55,56] However, as released GWAS statistics from this work did not contain beta coefficients, which is necessary for polygene score construction, we used summary statistics on the largest available component, i.e. based on fluid intelligence (max N samples = 108,818) from www.nealelab.is analysis extract of the UK Biobank.[33] In this case, fluid intelligence from the general population should capture any common genetic variants in the same fashion as Full Scale IQ and Verbal IQ, and thus should serve as a suitable proxy.

We built polygene scores using PRSice2 version 2.1.2 beta[57] under default conditions, i.e. using SNPs with an INFO score >0.90, r^2 of 0.10, and distance of 250 kbp, where r^2 was calculated on the target data (i.e. this cohort). We used pre-specified p-value cut-offs for SNPs for inclusion in the polygene score based on the p-value reported in the original GWAS that maximized previously reported prediction ability. For schizophrenia we used a p-value threshold of 0.05 (Extended Data Figure 5[23]), and for the UK Biobank Fluid Intelligence / IQ, we used 0.10 (Supplementary Table 2, largest explained variance in 2 out of 3 analyses[25]). For the schizophrenia polygene score, there were 80,496 SNPs after clumping, while for the IQ polygene score, there were 80,557 SNPs after clumping.

A priori power analyses and estimation of cohort specific parameter values

We conducted power analyses using simulations under a liability threshold model for our primary investigations using available sample sizes, known heritabilities, genetic correlations, and assumptions regarding the nature of the relationship between schizophrenia and subthreshold psychosis. All simulation results assume h2_g SZ = 0.46,[58] h2_g SZ (PRS) = 0.08,[23] h2_g IQ = 0.25,[25] h2_g (PRS) IQ = 0.04,[25] and r_g between SZ and IQ of -0.234.[25] In addition, in the absence of pre-existing literature estimates, we assumed h2_g subthreshold psychosis = 0.46 (based on h2_g SZ), and h2_g VIQ decline = 0.25 (based on h2_g IQ) **(Supplementary Table 4 and 5, and Supplementary Figure 5 and 6)**.

To estimate cohort specific parameters necessary for power analyses, we fit the observed data to a parametric likelihood based model based on the liability threshold model, with parameters as follows: schizophrenia prevalence; subthreshold psychosis prevalence; two shape parameters assuming the age distribution in the population following a beta binomial distribution; mean and SD for age at development of schizophrenia assuming a normal distribution; mean and SD for age at development of subthreshold psychosis assuming a normal distribution.

To explain the model, we considered a generative form, i.e. with population of individuals for study given the parameters above. Subsequently, we first simulated whether an individual would ever develop schizophrenia or subthreshold psychosis,

24

based on the prevalences of the two conditions (i.e. if the prevalence was 20%, then one would simulate phenotypes under a Bernoulli distribution with probability p=0.20). Next, independently, age was simulated, based on the shape parameters controlling the age distribution. Afterwards, age of diagnosis, conditional on ever developing the phenotype, was simulated, based on the parameters controlling the mean and SD age of development. From these underlying values for each simulated individual of the current age, whether they will ever develop schizophrenia or subthreshold psychosis, and the age at which they develop the phenotype, the present day phenotype of these simulated individuals could be determined.

Using constrained optimization, we obtained the parameters that maximized the likelihood of our real data under the above described model. To obtain a confidence interval for each parameter, we determined the maximum values of that parameter where twice the difference in log likelihood between the maximum likelihood estimation and that point was less than the chi-squared statistic with the appropriate number of degrees of freedom. Before applying the model to real data, we first simulated under the model to verify that we could recover parameter estimates on similar sized datasets, which confirmed the accuracy of the model (results not shown). We next generated parameter estimates on the real data (**Supplementary Table 4**). We used these parameter estimates and other literature derived estimates in the power analyses that were performed.

Regression analysis

We assessed relationships between PS_SZ and PS_IQ and binary phenotypes using logistic regression (SSD, subthreshold psychosis, VIQ decline), and linear regression for quantitative phenotypes (baseline FSIQ), adjusting for age, sex and the first five principal components from the imputed genotypes, with the principal components calculated using PLINK. All statistical tests in this manuscript are two-sided unless otherwise noted. r2 reported from linear regression is standard unadjusted r2 while from logistic regression is Nagelkerke r2.

Calculation of Positive Predictive Values (PPV)

We calculated PPV in the traditional way given binary phenotypes schizophrenia and ID and observed PS_SZ and PS_IQ among 22q11DS samples. We also estimated PPVs for the general population using known estimates of general population prevalence of schizophrenia and ID, as well as sensitivity and specificity values derived from our analysis in this 22q11DS sample.

Author contributions:

Drs. Vorstman, Davies, Fiksinski, and Bearden had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Drs. Davies and Fiksinski contributed equally to the study.

Study concept and design: AIM II writing group: Davies, Fiksinski, Breetvelt, Monfuega, Williams, Hooper, Bearden, Vorstman. *Acquisition of data:* Vorstman, Fiksinski, Duijff, Eliez, Schneider, Jalbrzikowski, Armando, Vicari, Shashi, Hooper, Chow, McDonald-McGinn, Vogels, Gothelf, Weinberger, van Amelsvoort, Kates, Antshel, Simon, Ousley, Swillen, Gur, Bearden, Bassett. *Analysis and interpretation of data:* AIM II writing group: Davies, Fiksinski, Breetvelt, Monfuega, Williams, Hooper, Bearden, Vorstman. *Critical revision of the manuscript for important intellectual content:* All authors. *Statistical analysis:* Davies, Breetvelt, Monfuega, Williams, Vorstman.

Potential competing interests disclosures:

Drs. Owen and van den Bree reported pharmaceuticals outside of the submitted work.

Dr Celso Arango has been a consultant to or has received honoraria or grants from

Acadia, Ambrosseti, Gedeon Richter, Janssen Cilag, Lundbeck, Otsuka, Roche, Sage,

Servier, Shire, Schering Plough, Sumitomo Dainippon Pharma, Sunovion and Takeda. Dr

D. Murphy has provided consultation to Roche. Dr. Hooper has provided consultation to

Novartis. Dr Ousley is a collaborator in a Biomarin Pharmaceutical study. None of the

other authors have financial disclosures.

Group authorship International 22q11.2 Brain and Behavior Consortium:

Michael P. Epstein	Department of Human Genetics, Emory University School of Medicine, Atlanta, GA, USA
Stephen T. Warren	Department of Human Genetics, Emory University School of Medicine, Atlanta, Georgia, United States of America.
Sinead Morrison	MRC Centre for Neuropsychiatric Genetics and Genomics, Division of Psychological Medicine and Clinical Neurosciences, School of Medicine, Cardiff University, Cardiff, UK
Samuel Chawner	MRC Centre for Neuropsychiatric Genetics and Genomics, Division of Psychological Medicine and Clinical Neurosciences, School of Medicine, Cardiff University, Cardiff, UK
Claudia Vingerhoets	School Mental Health & Neuroscience Maastricht University
Jeroen Breckpot	Center for Human Genetics, University Hospitals Leuven & Department of Human Genetics KU Leuven
Elfi Vergaelen	Center for Human Genetics, University Hospitals Leuven
Annick Vogels	Center for Human Genetics, University Hospitals Leuven & Department of Human Genetics KU Leuven
Stephen Monks	Department of Psychiatry, Royal College of Surgeons in Ireland, Education and Research Centre, Beaumont Hospital, Dublin, Ireland.
Sarah E. Prasad	Department of Psychiatry, Royal College of Surgeons in Ireland, Education and Research Centre, Beaumont Hospital, Dublin, Ireland.
Corrado Sandini	Developmental Imaging and Psychopathology, Department of Psychiatry, University of Geneva
Maude Schneider	Developmental Imaging and Psychopathology, Department of Psychiatry, University of Geneva
Johanna Maeder	Developmental Imaging and Psychopathology, Department of Psychiatry, University of Geneva
David Fraguas	Department of Child and Adolescent Psychiatry, Hospital General Universitario Gregorio Marañón, IiSGM, CIBERSAM, School of Medicine, Universidad Complutense, Madrid, Spain
Rens Evers	School Mental Health & Neuroscience Maastricht University
Flora Tassone	Department of Microbiology and Molecular Medicine, University of California Davis
Jaume Morey-Canyelles	Hospital Universitari Son Espases, Mallorca, Spain.
Opal Y. Ousley	Department of Psychiatry, Emory University School of Medicine
Kevin M. Antshel	Department of Psychology, Syracuse University
Wanda Fremont	Department of Psychiatry and Behavioral Sciences, SUNY Upstate Medical University
Rosemarie Fritsch	Department of Psychiatry, Hospital Clinico Universidad de Chile.
Claudia Ornstein	Department of Psychiatry, Hospital Clinico Universidad de Chile.

Eileen M. Daly	Department of Forensic and Neurodevelopmental Sciences, Institute of Psychiatry, Psychology & Neuroscience (IoPPN), King's College London
Gregory A. Costain	Department of Psychiatry, University of Toronto; Hospital for Sick Children, Toronto, Ontario, Canada
Erik Boot	Toronto General Hospital, University Health Network, Toronto, Ontario, Canada
Tracy Heung	Clinical Genetics Research Program, Centre for Addiction and Mental Health, Toronto, Ontario, Canada; Toronto General Hospital, University Health Network, Toronto, Ontario, Canada
T. Blaine Crowley	Division of Human Genetics and 22q and You Center, the Children's Hospital of Philadelphia, Philadelphia, PA, USA
Elaine H. Zackai	Division of Human Genetics and 22q and You Center, the Children's Hospital of Philadelphia and Department of Pediatrics, Perelman School of Medicine of the University of Pennsylvania, Philadelphia, PA, USA
Monica E. Calkins	Department of Psychiatry and Lifespan Brain Institute, Penn Medicine - CHOP, University of Pennsylvania
Ruben C. Gur	Department of Psychiatry and Lifespan Brain Institute, Penn Medicine - CHOP, University of Pennsylvania
Kathryn L. McCabe	University of Newcastle, University of California Davis
Tiffany Busa	Departement de Genetique Medicale Hôpital d'Enfants de la Timone, APHM, Marseille, France
Kelly Schoch	Department of Pediatrics, Division of Medical Genetics, Duke University School of Medicine
Maria Pontillo	Child and Adolescence Neuropsychiatry Unit, Department of Neuroscience, IRCSS Bambino Gesù Children's Hospital of Rome
Sasja N. Duijff	Department of Pediatrics University Medical Center Utrecht
René S. Kahn	Brain Center Rudolf Magnus, University Medical Center Utrecht, Netherlands, and Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, New York.
Michiel Houben	Department of Pediatrics University Medical Center Utrecht
Leila Kushan	Departments of Psychiatry and Biobehavioral Sciences and Psychology, Semel Institute for Neuroscience and Human Behavior. University of California, Los Angeles
Maria Jalbrzikowski	Department of Psychiatry, University of Pittsburgh School of Medicine
Miri Carmel	Sackler Faculty of Medicine, Tel Aviv University; Felsenstein Medical Research Center Petach Tikva
Ehud Mekori-Domachevsky	The Child Psychiatry Division, Sheba Medical Center, Tel Hashomer; Sackler Faculty of Medicine, Tel Aviv University
Elena Michaelovsky	Sackler Faculty of Medicine, Tel Aviv University; Felsenstein Medical Research Center Petach Tikva
Ronnie Weinberger	The Child Psychiatry Division, Sheba Medical Center, Tel Hashomer

Data availability statement
The datasets (raw data) generated during and/or analysed during the current study are being made available through the NIMH Data Archive (NDA) repository, [accession number forthcoming].

Code availability statement
Bespoke analysis code for analyses downstream of genotype generation is available https://github.com/rwdavies/IBBC_Aim2_22Q11DS.

List of participating institutions

Institution	City, Country
Aix-Marseille University	Marseille, France
Albert Einstein College of Medicine	New York, USA
Bambino Gesu Ospedale	Rome, Italy
Cardiff University	Cardiff, Wales
Center for Addiction and Mental Health, Toronto General Hospital and the University of Toronto	Toronto, Canada
Children's Hospital of Philadelphia	Philadelphia, USA
Duke University	Durham, USA
Emory University	Atlanta, USA
Hospital Son Espases	Palma, Spain
Hospital Universitario La Paz	Madrid, Spain
Katholieke University, Leuven	Leuven, Belgium
King's College London	London, England
Maastricht University	Maastricht, the Netherlands
Royal College of Surgeons Ireland	Dublin, Ireland
SUNY Upstate Medical University	Syracuse, USA
Syracuse University	Syracuse, USA
Tel Aviv University and Sheba Medical Center	Tel Aviv, Israel
The Hospital for Sick Children	Toronto, Canada
Universidad del Desarrollo	Santiago, Chile
University Medical Center Utrecht	Utrecht, the Netherlands
University of New Castle	Newcastle, Australia
University of North Carolina Charlotte	Chapel Hill, USA
University of California Davis	Sacramento, USA
University of California Los Angeles	Los Angeles, USA
University of Geneva	Geneva, Switzerland
Perelman School of Medicine of the University of Pennsylvania	Philadelphia, USA

Literature Cited

1. Murphy, K.C., Jones, L.A. & Owen, M.J. High rates of schizophrenia in adults with velo-cardio-facial syndrome. *Archives of general psychiatry* **56**, 940-945 (1999).
2. Bassett, A.S. & Chow, E.W. 22q11 deletion syndrome: a genetic subtype of schizophrenia. *Biol.Psychiatry* **46**, 882-891 (1999).
3. Sommer, I.E., *et al.* Early interventions in risk groups for schizophrenia: what are we waiting for? *NPJ Schizophr* **2**, 16003 (2016).
4. Reichenberg, A., *et al.* Static and dynamic cognitive deficits in childhood preceding adult schizophrenia: a 30-year study. *The American journal of psychiatry* **167**, 160-169 (2010).
5. Bearden, C.E., *et al.* A prospective cohort study of childhood behavioral deviance and language abnormalities as predictors of adult schizophrenia. *Schizophrenia bulletin* **26**, 395-410 (2000).
6. Rosso, I.M., *et al.* Childhood neuromotor dysfunction in schizophrenia patients and their unaffected siblings: a prospective cohort study. *Schizophrenia bulletin* **26**, 367-378 (2000).
7. Walker, E.F., Grimes, K.E., Davis, D.M. & Smith, A.J. Childhood precursors of schizophrenia: facial expressions of emotion. *The American journal of psychiatry* **150**, 1654-1660 (1993).
8. Dickson, H., Laurens, K.R., Cullen, A.E. & Hodgins, S. Meta-analyses of cognitive and motor function in youth aged 16 years and younger who subsequently develop schizophrenia. *Psychological medicine* **42**, 743-755 (2012).
9. Woodberry, K.A., Giuliano, A.J. & Seidman, L.J. Premorbid IQ in schizophrenia: a meta-analytic review. *The American journal of psychiatry* **165**, 579-587 (2008).
10. Mollon, J., David, A.S., Zammit, S., Lewis, G. & Reichenberg, A. Course of Cognitive Development From Infancy to Early Adulthood in the Psychosis Spectrum. *JAMA psychiatry* **75**, 270-279 (2018).
11. Zammit, S., *et al.* A longitudinal study of premorbid IQ Score and risk of developing schizophrenia, bipolar disorder, severe depression, and other nonaffective psychoses. *Archives of general psychiatry* **61**, 354-360 (2004).
12. Khandaker, G.M., Barnett, J.H., White, I.R. & Jones, P.B. A quantitative meta-analysis of population-based studies of premorbid intelligence and schizophrenia. *Schizophr Res* **132**, 220-227 (2011).
13. Kahn, R.S. & Keefe, R.S. Schizophrenia Is a Cognitive Illness: Time for a Change in Focus. *JAMA psychiatry* (2013).
14. MacCabe, J.H., *et al.* Decline in cognitive performance between ages 13 and 18 years and the risk for psychosis in adulthood: a Swedish longitudinal cohort study in males. *JAMA psychiatry* **70**, 261-270 (2013).
15. Meier, M.H., *et al.* Neuropsychological decline in schizophrenia from the premorbid to the postonset period: evidence from a population-

representative longitudinal study. *The American journal of psychiatry* **171**, 91-101 (2014).

16. Lin, A.*, et al.* Neurocognitive predictors of functional outcome two to 13 years after identification as ultra-high risk for psychosis. *Schizophr Res* **132**, 1-7 (2011).

17. Kaymaz, N.*, et al.* Do subthreshold psychotic experiences predict clinical outcomes in unselected non-help-seeking population-based samples? A systematic review and meta-analysis, enriched with new results. *Psychological medicine* **42**, 2239-2253 (2012).

18. Poulton, R.*, et al.* Children's self-reported psychotic symptoms and adult schizophreniform disorder: a 15-year longitudinal study. *Archives of general psychiatry* **57**, 1053-1058 (2000).

19. Insel, T.*, et al.* Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *The American journal of psychiatry* **167**, 748-751 (2010).

20. Gur, R.E.*, et al.* A neurogenetic model for the study of schizophrenia spectrum disorders: the International 22q11.2 Deletion Syndrome Brain Behavior Consortium. *Molecular psychiatry* (2017).

21. Vorstman, J.A.*, et al.* Cognitive decline preceding the onset of psychosis in patients with 22q11.2 deletion syndrome. *JAMA psychiatry* **72**, 377-385 (2015).

22. Pardinas, A.F.*, et al.* Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nature genetics* **50**, 381-389 (2018).

23. Schizophrenia Working Group of the Psychiatric Genomics, C. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421-427 (2014).

24. Wray, N.R.*, et al.* Research review: Polygenic methods and their application to psychiatric traits. *Journal of child psychology and psychiatry, and allied disciplines* **55**, 1068-1087 (2014).

25. Davies, G.*, et al.* Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. *Nat Commun* **9**, 2098 (2018).

26. Rietveld, C.A.*, et al.* GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science* **340**, 1467-1471 (2013).

27. Sugrue, L.P. & Desikan, R.S. What Are Polygenic Scores and Why Are They Important? *JAMA* (2019).

28. Fullerton, J.M. & Nurnberger, J.I. Polygenic risk scores in psychiatry: Will they be useful for clinicians? *F1000Res* **8**(2019).

29. Torkamani, A., Wineinger, N.E. & Topol, E.J. The personal and clinical utility of polygenic risk scores. *Nature reviews. Genetics* **19**, 581-590 (2018).

30. Kuchenbaecker, K.B.*, et al.* Evaluation of Polygenic Risk Scores for Breast and Ovarian Cancer Risk Prediction in BRCA1 and BRCA2 Mutation Carriers. *J Natl Cancer Inst* **109**(2017).

31. Cleynen, I.*, et al.* Genetic contributors to risk of schizophrenia in the presence of a 22q11.2 deletion. *Molecular psychiatry* (2020).

32. Marder, S.R. & Cannon, T.D. Schizophrenia. *The New England journal of medicine* **381**, 1753-1761 (2019).
33. Bycroft, C.*, et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203-209 (2018).
34. Chelune, G.J., Naugle, R.I., Lüders, H., Sedlak, J. & Awad, I.A. Individual Change After Epilepsy Surgery: Practice Effects and Base-Rate Information. *Neuropsychology.* **7**, 41-52 (1993).
35. Klaassen, P.*, et al.* Explaining the variable penetrance of CNVs: Parental intelligence modulates expression of intellectual impairment caused by the 22q11.2 deletion. *American journal of medical genetics. Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics* **171**, 790-796 (2016).
36. Fuller, R.*, et al.* Longitudinal assessment of premorbid cognitive functioning in patients with schizophrenia through examination of standardized scholastic test performance. *The American journal of psychiatry* **159**, 1183-1189 (2002).
37. van Oel, C.J., Sitskoorn, M.M., Cremer, M.P. & Kahn, R.S. School performance as a premorbid marker for schizophrenia: a twin study. *Schizophrenia bulletin* **28**, 401-414 (2002).
38. Duijff, S.N.*, et al.* Cognitive development in children with 22q11.2 deletion syndrome. *The British journal of psychiatry : the journal of mental science* **200**, 462-468 (2012).
39. Chawner, S.*, et al.* Childhood cognitive development in 22q11.2 deletion syndrome: case-control study. *The British journal of psychiatry : the journal of mental science* **211**, 223-230 (2017).
40. Trevethan, R. Sensitivity, Specificity, and Predictive Values: Foundations, Pliabilities, and Pitfalls in Research and Practice. *Front Public Health* **5**, 307 (2017).
41. Bergen, S.E.*, et al.* Joint Contributions of Rare Copy Number Variants and Common SNPs to Risk for Schizophrenia. *The American journal of psychiatry* **176**, 29-35 (2019).
42. Tansey, K.E.*, et al.* Common alleles contribute to schizophrenia in CNV carriers. *Molecular psychiatry* **21**, 1085-1089 (2016).
43. Lecarpentier, J.*, et al.* Prediction of Breast and Prostate Cancer Risks in Male BRCA1 and BRCA2 Mutation Carriers Using Polygenic Risk Scores. *J Clin Oncol* **35**, 2240-2250 (2017).
44. Gibson, G. On the utilization of polygenic risk scores for therapeutic targeting. *PLoS genetics* **15**, e1008060 (2019).
45. Damask, A.*, et al.* Patients With High Genome-Wide Polygenic Risk Scores for Coronary Artery Disease May Receive Greater Clinical Benefit From Alirocumab Treatment in the ODYSSEY OUTCOMES Trial. *Circulation* **141**, 624-636 (2020).
46. Perkins, D.O.*, et al.* Polygenic Risk Score Contribution to Psychosis Prediction in a Target Population of Persons at Clinical High Risk. *The American journal of psychiatry* **177**, 155-163 (2020).

47. Fiksinski, A.M.*, et al.* Understanding the pediatric psychiatric phenotype of 22q11.2 deletion syndrome. *American journal of medical genetics. Part A* (2018).

48. Martin, A.R.*, et al.* Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. *American journal of human genetics* **100**, 635-649 (2017).

49. Vorstman, J.A.*, et al.* MLPA: a rapid, reliable, and sensitive method for detection and analysis of abnormalities of 22q. *Hum Mutat* **27**, 814-821 (2006).

50. Miller, T.J.*, et al.* Prodromal assessment with the structured interview for prodromal syndromes and the scale of prodromal symptoms: predictive validity, interrater reliability, and training to reliability. *Schizophrenia bulletin* **29**, 703-715 (2003).

51. Yung, A.R.*, et al.* Mapping the onset of psychosis: the Comprehensive Assessment of At-Risk Mental States. *Aust.N.Z.J.Psychiatry* **39**, 964-971 (2005).

52. Kaufman, J.*, et al.* Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. *Journal of the American Academy of Child and Adolescent Psychiatry* **36**, 980-988 (1997).

53. Insel, T.R. Rethinking schizophrenia. *Nature* **468**, 187-193 (2010).

54. Purcell, S.*, et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am.J.Hum.Genet.* **81**, 559-575 (2007).

55. Deary, I.J., Johnson, W. & Houlihan, L.M. Genetic foundations of human intelligence. *Human genetics* **126**, 215-232 (2009).

56. Spearman, C. "General Intelligence," objectively determined and measured. *American Journal of Psychology* **15**, 201-293 (1904).

57. Euesden, J., Lewis, C.M. & O'Reilly, P.F. PRSice: Polygenic Risk Score software. *Bioinformatics* **31**, 1466-1468 (2015).

58. Gazal, S.*, et al.* Linkage disequilibrium-dependent architecture of human complex traits shows action of negative selection. *Nature genetics* **49**, 1421-1427 (2017).

Supplementary Material

Table of Contents

1. Supplementary Note 1: Post-hoc investigations into the relationship between subthreshold psychosis and PS_SZ

We observed a nominally significant association between subthreshold psychosis and PS_SZ. This observation did not meet multiple testing criterion. Post-hoc, we performed three analyses to explore the observed association between subthreshold psychosis and PS_SZ

Methods

Effect of future SSD cases as a source of confounding between subthreshold psychosis and PS_SZ

We modelled a scenario whereby the PS_SZ signal would be driven by the presence of individuals with future, as of yet undiagnosed SSD in the subthreshold psychosis group. In essence, we estimated in this scenario what proportion of such future SSD cases would be required to explain the observed PS_SZ in the subthreshold psychosis group

Genetically correlated traits as a source of confounding between subthreshold psychosis and PS_SZ

We examined whether the observed PS_SZ results in the subthreshold psychosis group could originate from increased rates of other psychiatric phenotypes that are genetically correlated with schizophrenia. Available IBBC data allowed us to analyse this possibility for comorbid mood disorders. Underlying assumptions for our mediation analysis were based on extrapolations of the IBBC data and include increased rates of (future) SSD

(~40%) and mood disorder (49%) in the subthreshold psychosis group, compared to ~17% rates for both phenotypes in controls.

Quantitative measure of subthreshold psychosis as additional evidence for relationship between subthreshold psychosis and PS_SZ

In a subset of 347 of 962 individuals with a well-defined phenotype and imputed genotype data, we were able to obtain an integer-coded measure of subthreshold psychosis from the Structured Interview for Prodromal Syndromes (SIPS). We first generated a transformation from the integer coded, non-normally distributed quantitative SIPS score by fitting an exponential distribution using the least square estimate, yielding a transformation, in R, of "qnorm(pexp(q = x + 0.5, rate = 0.2238))", where x is the original integer coded SIPS score (**Supplementary Figure 7**). This yields a more approximately normally distributed value.

We assessed power to detect an association between the quantitative SIPS based phenotype and PS_SZ using simulations. Using the same assumptions listed before regarding heritabilities and predictive accuracies of polygenic scores, we first simulated an underlying total liability (genetic and environmental) for the quantitative subthreshold psychosis. As before, this total liability becomes binary under a liability threshold model, giving us the binary definition of subthreshold psychosis. In addition, using the continuous total liability, we generated an integer coded value (representing a simulated SIPS score) as "round(qexp(pnorm(Y_sub), rate = 0.2238))". We then re-

transformed this to a continuous value using it's inverse "qnorm(pexp(q = x + 0.5, rate = 0.2238))", and from this, could calculate power for detecting an association between the quantitative subthreshold psychosis and PS_SZ, with or without conditioning on the binary phenotype (**Supplementary Figure 8**).

Results

First, given that some fraction of individuals with subthreshold psychosis will eventually develop SSD, we modelled what proportion would need to develop SSD to be consistent with our findings. The observed levels of PS_SZ are consistent with a scenario in which 86% (95% CI 56 - 100%) of individuals with subthreshold psychosis would in fact represent future SSD patients who were not yet identified as such at the time of the assessment. This is a proportion inconsistent with known rates of SSD in 22q11DS (see **Supplementary Figure 9**)[1], rendering it unlikely that our result is driven by "future" SSD cases.

Second, we examined potential confounding by psychiatric comorbidity genetically correlated with SSD. In this sample, the rate of comorbid mood disorders in the subthreshold psychosis group was 29.2%, versus 22.7% in the merged controls (**Table 1**). Results from the mediation analysis indicated a lack of attenuation through the mood disorder phenotype (effect size of PS_SZ in model without mood disorder is 0.239, p = 0.025; with mood disorder 0.250, p = 0.021) (**Supplementary Table 6**), indicating that the observed increased PS_SZ in subthreshold psychosis is not readily explained by the higher rate of mood disorders in this group.

Third, we explored the correlation between a quantitative measure of subthreshold psychosis severity obtained in N = 347 individuals and PS_SZ as additional evidence. However, when adjusting for the previous binary indicator of subthreshold psychosis versus control, the association between the transformed quantitative SIPS phenotype and PS_SZ was not significant (N = 347, p = 0.77, r^2 = 0.0001; **Supplementary Figure 10**).

Interpretation

First, we showed that undiagnosed "future" cases are insufficient to explain the observed signal.

Second, we assessed the effect of comorbid mood disorders in the sample, given the reported genetic correlation between mood disorders and schizophrenia,[2] and again did not find evidence that explained the signal.

Finally, we examined residual quantitative variation in subthreshold psychosis. The association between this transformed quantitative variable and PS_SZ was not significant when adjusting for the previous binary indicator of these phenotypes. However, it is worth noting that a priori power for this analysis was limited and dependent on strong assumptions. Interestingly, studies on genetic correlations between subthreshold psychotic symptoms and PS_SZ in the general population reported to date are conflicting[3-5], impeding definite evidence in this regard.

We conclude that our findings tentatively suggest a genetic correlation between subthreshold psychotic symptoms and schizophrenia in 22q11DS, but that further studies are required to provide more certainty in this regard.

2. Supplementary Note 2: Inverse probability weight

We hypothesized that for putative controls, the probability of not developing SSD (i.e., remaining a control), is following the inverse curve of the cumulative incidence for schizophrenia in 22q11DS.

Methods

In order to test this hypothesis, we used a linear regression model to investigate the function of PS_SZ and age in putative controls, correcting for sex and the first two components of the PCA. Subsequently, we investigated whether the PS_SZ analysis comparing 22q11DS cases with SSD to merged controls changed when taking into account the probability for each putative control to develop SSD. To this end, we used an inversed probability weight, which was based on our hypothesized function of the number of as of yet unidentified future SSD cases. Based on this model we assumed that number of false negatives (i.e., putative controls that will eventually develop SSD) remains stable at around 20% until age 10 years and subsequently decreases throughout adolescence until early adulthood (**Supplementary Figure 11**).

Results

As expected, we observed a small negative linear correlation between age and PS_SZ in the putative control group (r^2 = -0.2, p =0.048, **Supplementary Figure 12**). PR_SZ associations of schizophrenia and subthreshold psychosis versus controls were not

altered significantly when incorporating an inverse probability weight based on the

observed correlation (data not shown).

3. Supplementary Note 3: Examining the potential impact of ascertainment bias on observed genetic effect

Given the possibility of ascertainment biases in our cohort, we performed simulations to examine the potential influence of variations in phenotypic prevalence (due to ascertainment) on the observed genetic effects in our study.

Methods

In the simulated dataset, let:

i) The prevalence of schizophrenia in the entire population be 1%

ii) The fraction of variance explained by a measurable PS for schizophrenia be 5%

iii) The prevalence of schizophrenia given you have 22q11 be 25%

iv) The frequency of 22q11DS be 1/3000

For some very large N, sample one individual under a liability threshold model as having:

i) A measurable PS component of liability sampled from a normal mean 0 variance 0.05

ii) All other risk sampled from a normal having mean 0 variance 0.95

iii) From this we construct a genetic liability as the sums of these two contributions

iv) We further simulate with 1/3000 frequency whether someone has 22q11DS, and if they do, add to the liability qnorm(1 - 0.01) - qnorm(1 - 0.25) i.e. add to the liability an additive component sufficient to make the baseline risk 25%

Across 32 repetitions we drew 10,000 sampled individuals from an initial sampling of

100,000,000, and then tested scenarios with

i) Sampling at random from those with 22q11DS

ii) Sampling 90% at random, 10% sampled conditional on having SCZ

iii) Sampling 75% at random, 25% sampled conditional on having SCZ

Results

On average across these, we calculated prevalence, PPVs in each quantile, and OR of

having a PS above or below the median PS:

i)   Prevalence = 26%, Quintile PPVs = 16, 22, 25, 30, 37, OR = 1.91

ii)   Prevalence = 34%, Quintile PPVs = 22, 29, 33, 38, 46, OR = 1.89

iii) Prevalence = 45%, Quintile PPVs = 32, 40, 45, 51, 59, OR = 1.90

Interpretation

Assuming these simulation settings, intended to be realistic with respect to our own

dataset and the underlying assumptions, we observe that while prevalence estimates do

vary, and that this affects the observed within-quintile PPV, the effect sizes are the same

between groups, as exemplified by the OR using a binary cutoff of above or below the

median polygenic score. These findings indicate that polygenic scores can add

meaningfully to risk stratification in the setting of a high risk CNV, and that this

observation is robust despite the effect on ascertainment-related biases in phenotypic

prevalences.

4. Supplementary Tables 1-6

	Schizophrenia spectrum diagnosis (SSD)	Merged controls (all ages)
$90 < PS^{ile}$	OR = 1.44 [0.88, 2.37] PPV = 0.325 [0.222, 0.428] N+ = 26, N- = 181	N+ = 54, N- = 543
$75 < PS^{ile}$	OR = 1.62 [1.14, 2.31] PPV = 0.332 [0.265, 0.398] N+ = 64, N- = 143	N+ = 129, N- = 468
$50 < PS^{ile}$	OR = 1.91 [1.38, 2.64] PPV = [0.274, 0.366] N+ = 126, N- = 81	N+ = 268, N- = 329
$PS^{ile} < 50$	OR = 0.52 [0.38, 0.72] PPV = 0.198 [0.159, 0.236] N+ = 81, N- = 126	N+ = 329, N- = 268
$PS^{ile} < 25$	OR = 0.41 [0.27, 0.62] PPV = 0.147 [0.099, 0.195] N+ = 31, N- = 176	N+ = 180, N- = 417
$PS^{ile} < 10$	OR = 0.26 [0.12, 0.55] PPV = 0.091 [0.031, 0.151] N+ = 8, N- = 199	N+ = 80, N- = 517

Supplementary Table 1. OR for schizophrenia based on polygenic score cutoffs. Results show for a given binary cutoff based on Polygenic Score percentile, how many 22q11.2DS individuals fall above or below that cutoff, stratified by having SSD, or being a control (regardless of age). ORs and PPVs are given for SSD against merged controls. Prevalence of SSD (observed) is 26% (versus controls).

FSIQ Polygenic Score cutoff (percentile)	22q11.2DS with ID	22q11.2DS without ID
PS^{ile} < 10	OR = 2.64 [1.59, 4.4 PPV = 0.629 [0.515, 0.742] , N+ = 44, N- = 246	N+ = 26, N- = 384
PS^{ile} < 25	OR = 2.07 [1.47, 2.93] PPV = 0.549 [0.475, 0.622] N+ = 96, N- = 194	N+ = 79, N- = 331
PS^{ile} < 50	OR = 1.85 [1.37, 2.51] PPV = 0.489, [0.436, 0.541] N+ = 171, N- = 119	N+ = 179, N- = 231
50 < PS^{ile}	OR = 0.54 [0.4, 0.73] PPV = 0.34 [0.29, 0.39] N+ = 119, N- = 171	N+ = 231, N- = 179
75 < PS^{ile}	OR = 0.49 [0.34, 0.71] PPV = 0.291 [0.224, 0.359] N+ = 51, N- = 239	N+ = 124, N- = 286
90 < PS^{ile}	OR = 0.42 [0.24, 0.74] PPV = 0.243, [0.142, 0.343] N+ = 17, N- = 273	N+ = 53, N- = 357

Supplementary Table 2. OR and PPV for ID based on polygenic score cut-offs. Results show for a given binary cut-off based on Polygenic Score percentile, how many 22q11.2DS individuals fall above or below that cut-off, stratified by having ID or not having ID. Odds-ratios and PPVs are shown for each percentile cut-off. ID is defined as IQ < 70. Overall prevalence of ID is 41%.

		Mean Age (SD)	Sex (%M)	SSD	Control	Putative ctrl	Sub-threshold	No pheno data
Included	962	24.0 (12.4)	48.6%	207 (21.5%)	215 (22.3%)	382 (39.7%)	158 (16.4%)	0
Excluded	824	18.2 (9.1)	48.5%	123 (14.9%	79 (9.6%)	468 (56.8%)	110 (13.3%)	11 (1.3%)

Supplementary Table 3. An overview of demographic differences between included and excluded individuals in the study. Note that these differences reflect the history of the IBBC recruitment strategy. In "phase 1" submission of DNA from individuals who were either (schizophrenia spectrum) case or true control (age >25) was encouraged. There are more Affymetrix data available from this "phase 1", because by the time the second wave started, the WGS effort was up and running. This "phase 2" also included individuals who did not directly qualify as either case or definitive control. The main reason for exclusion for the current study was lack of availability of Affymetrix data. Therefore, as a result of the said prioritization of phase 1 (schizophrenia cases and definitive controls), the mean age of subjects with available Affymetrix data is also higher compared to those without Affymetrix data (enriched in phase 2). The age difference occurred because the onset of schizophrenia is generally after age 18 years, and true controls were defined as only those without psychosis and older than 25 years. In addition, given the on average lower age range in the individuals with no available Affymetrix data (hence: not included in this study), it is expected that the proportion of putative controls is higher in the excluded samples.

	Estimate	CI Lower Bound	CI Upper Bound
K_SZ	0.45	0.37	0.56
K_subthreshold psychosis	0.32	0.28	0.4
age_shape1	1.63	1.48	1.79
age_shape2	3.58	3.23	3.95
SZ_mean_age	23.05	19.64	27.98
SZ_sd_age	10.33	7.26	15.12
Subthreshold psychosis_mean_age	9.94	8.82	11.89
Subthreshold psychosis _sd_age	1.69	0.76	6.48

Supplementary Table 4. Parameter estimates for model that was used to inform power calculations in this study.

Dependent variable	IV	Power (alpha = 0.05)*	Relevant Supplementary Figure
SSD		0.997 [r_g = 1]	5
Subthreshold psychosis		0.062 [r_g = 0], 0.974 [r_g = 0.95]	5
Baseline FSIQ	PS_SZ	0.32 [r_g = 1]	6
VIQ decline		0.058 [r_g = 0], 1 [r_g = 0.8]	6
SSD		0.189 [r_g = 1]	5
Subthreshold psychosis		0.048 [r_g = 0], 0.867 [r_g = 0.95]	5
Baseline FSIQ	PS_IQ	1 [r_g = 1]	6
VIQ decline		0.048 [r_g = 0], 0.996 [r_g = 0.8]	6

Supplementary Table 5. Power analyses for primary analyses regarding genetic relationships between dependent variables (phenotypes) and independent variables (polygenic score). r_g is given between dependent variable and either schizophrenia (first four rows), or IQ (last four rows). Values of r_g between schizophrenia and IQ are fixed at -0.234, while otherwise, conditional on this, we report power for minimum and maximum possible genetic correlation between dependent variable and independent variable. IV = Independent Variable.

Purpose	Regression (bold = what in p, effect columns)	N	effect size	p
1 IV and mediator	mediator ~ **IV** + covars	943	-0.405	0.001
1* IV and mediator *(no schizophrenia cases)*	mediator ~ **IV** + covars	725	-0.221	0.202
2 IV and DV *(same as original)*	DV ~ **IV** + covars	755	0.239	0.025
3 mediator and DV	DV ~ **mediator** + covars	725	1.886	0.067
4 IV, DV and mediator	DV ~ **IV** + mediator + covars	725	0.250	0.021

Supplementary Table 6. Mediation analysis between subthreshold psychosis, mood disorders, and PS_SZ. IV = Independent Variable, DV = Dependent Variable.

5. Supplementary Figures 1-12
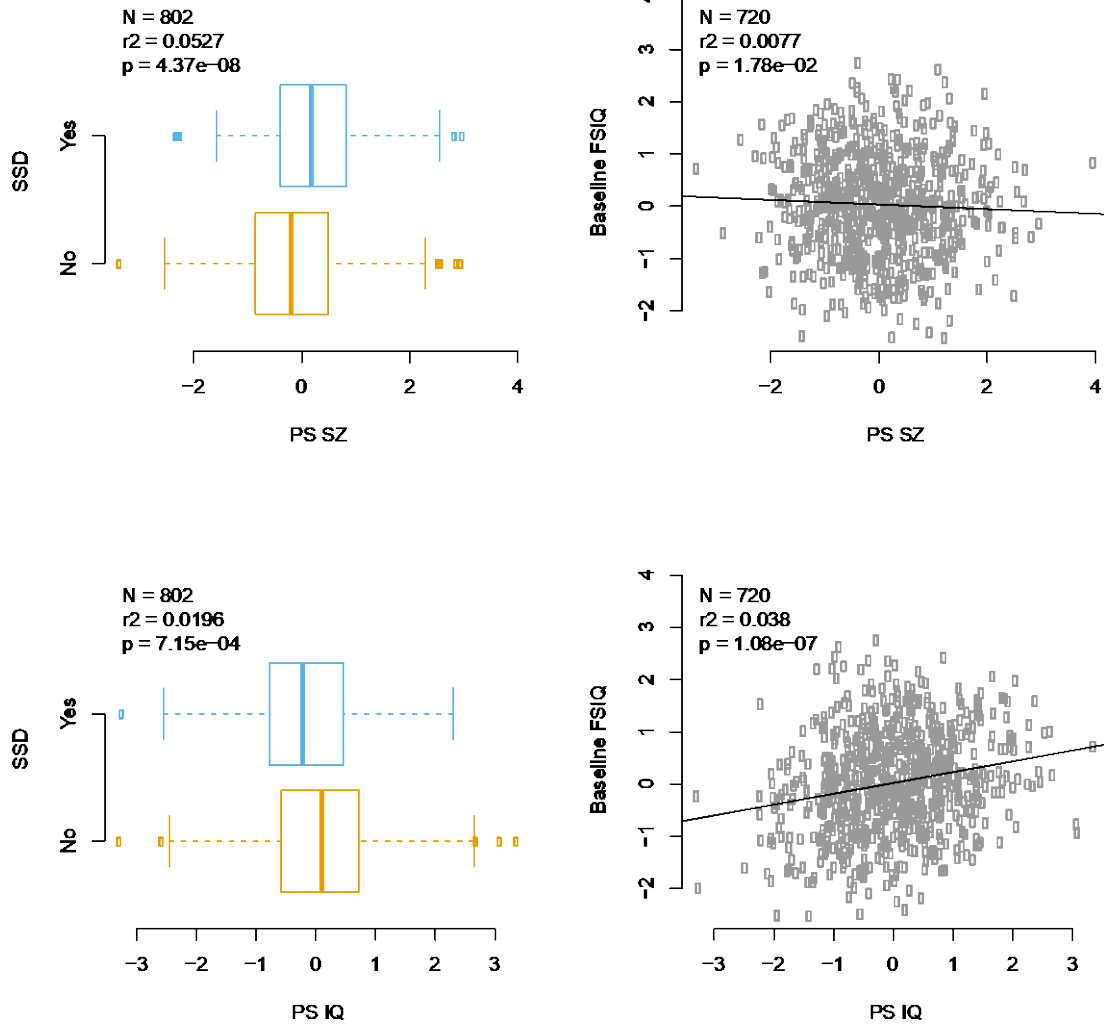
Supplementary Figure 1. Polygenic scores on schizophrenia and intellectual ability. Shown are results with previously identified relationships for phenotypes (columns; SSD and baseline FSIQ) and polygenic scores (rows; PS_SZ and PS_IQ). Results are visualized as either box-plots (binary phenotype) or scatterplot (quantitative phenotype). Box-plots highlight median, interquartile range and 5th/95th percentiles. Key regression results are highlighted in the plot (N total, marginal r^2, and p-value). Colors are to facilitate visualization only.
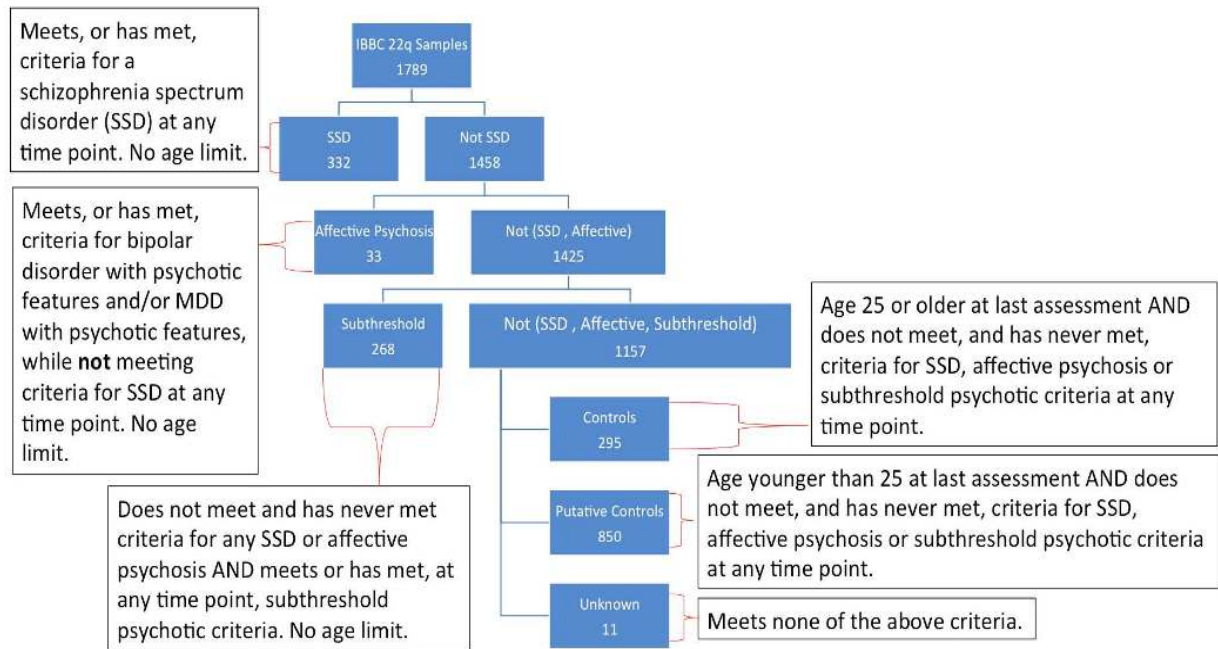
Supplementary Figure 2. Flowchart of IBBC cohort (full cohort) outlining the criteria used to assign IBBC subjects into different diagnostic classes regarding schizophrenia spectrum disorder (SSD) and related phenotypes.
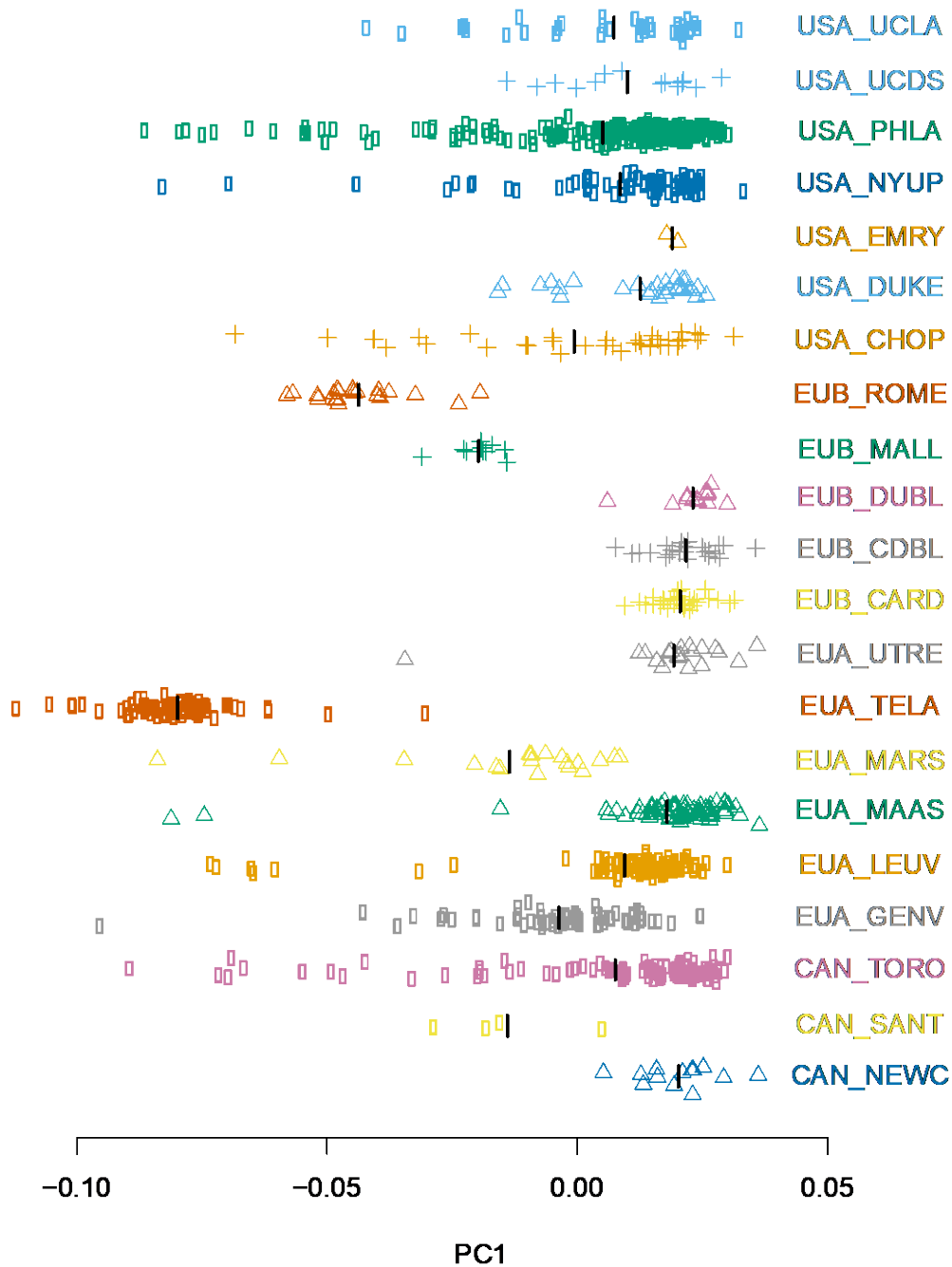
Supplementary Figure 3. Principal component 1 as function of study site. X-axis denotes value per-individual on PC1, while Y-axis is arbitrary to separate study sites plus jitter. Different sites are separated vertically and are grouped together by colour and plot icon. Black vertical bar indicates per study site average.
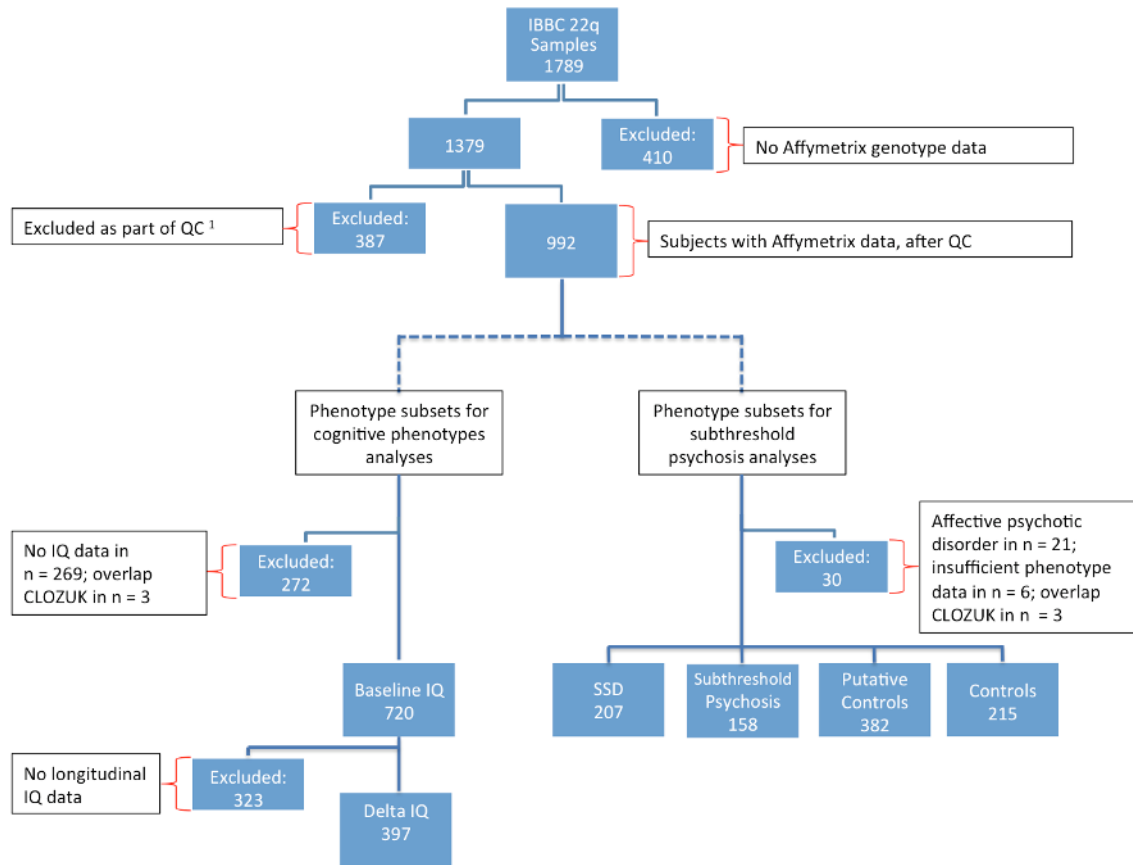
Supplementary Figure 4. Flowchart of subjects of IBBC cohort, outlining the different phenotypic subsets for the current study. [1.] Arrays excluded for sex coding reasons (5); missingness (84); IBD analysis (174); PCA (124).
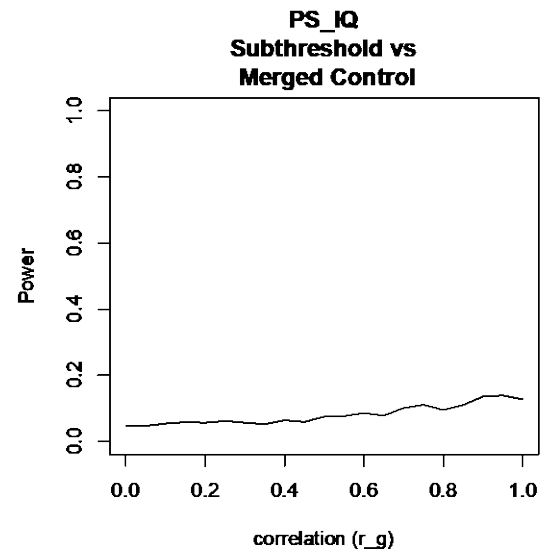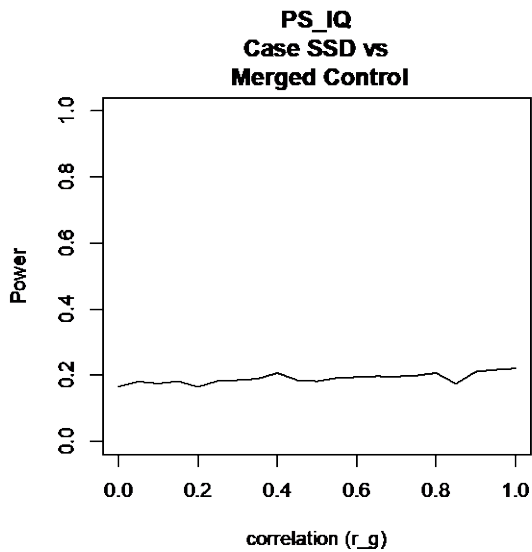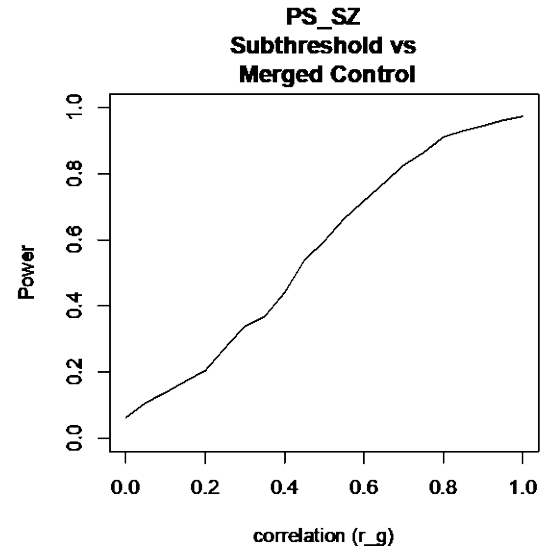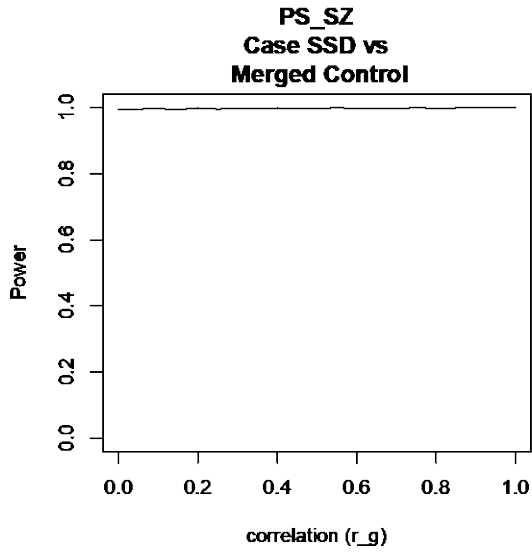
Supplementary Figure 5. Power to differentiate SSD status given genetic correlation. Shown are power at alpha=0.05 when comparing groups as specified in the plot subtitles for their difference in polygene score as specified in the title, given genetic correlation between subthreshold psychosis and SSD.
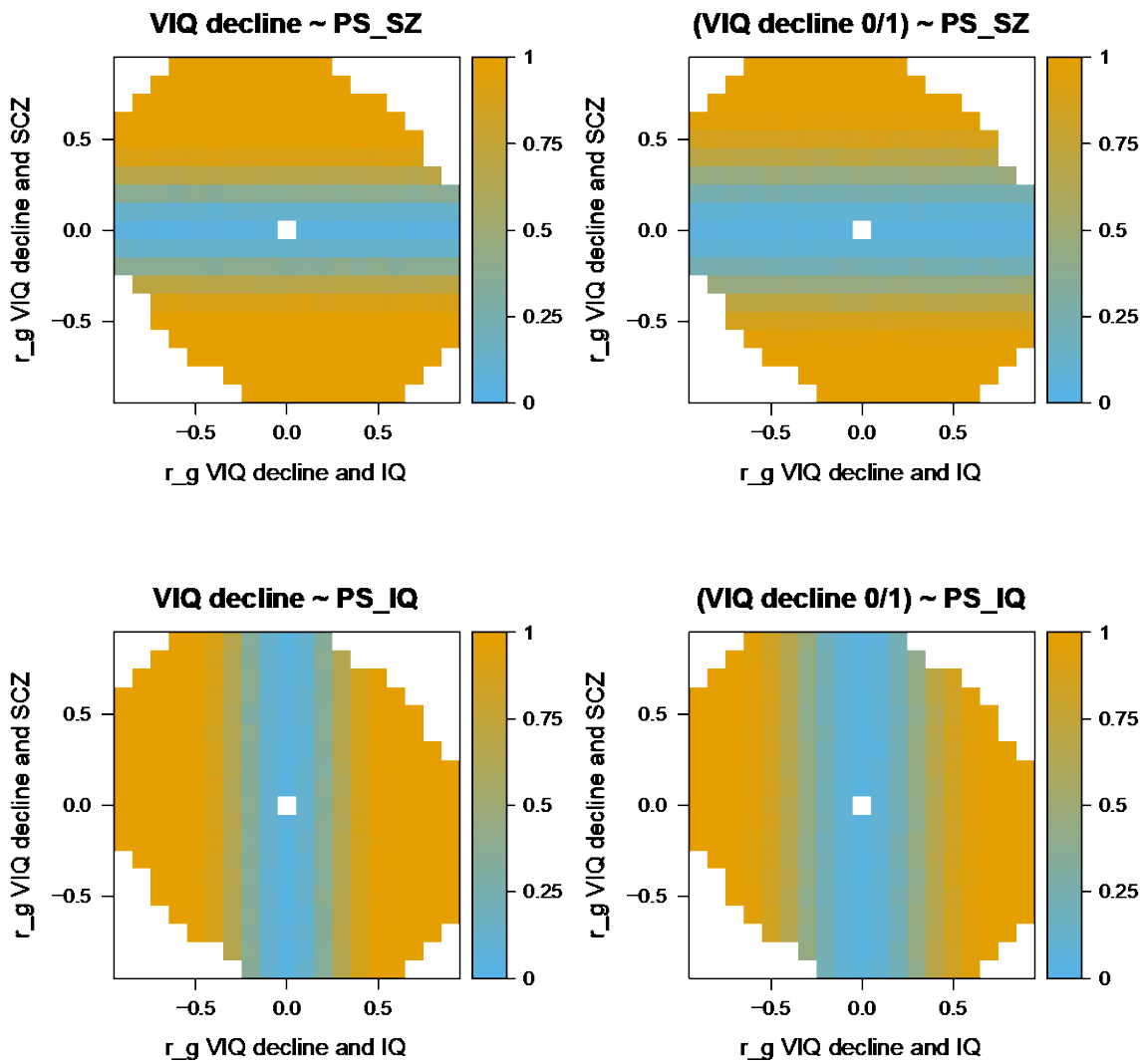
Supplementary Figure 6. Power to differentiate VIQ decline given genetic correlations. Shown are power at alpha=0.05 when regressing continuous or binary VIQ decline against PS_SZ or PS_IQ, shown as a function of both the genetic correlation between VIQ decline and IQ, as well as between VIQ decline and SSD.
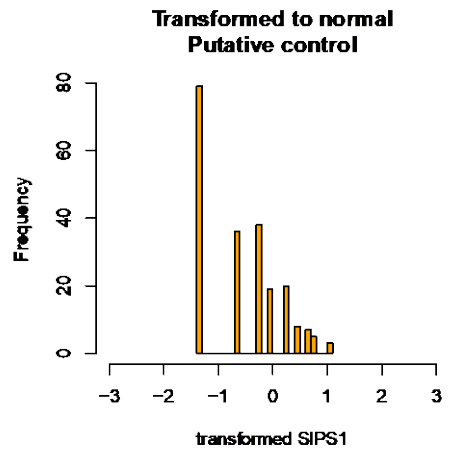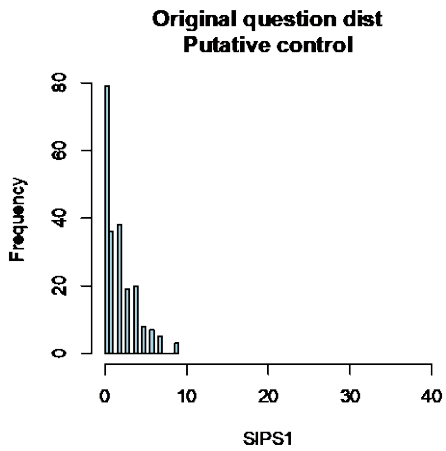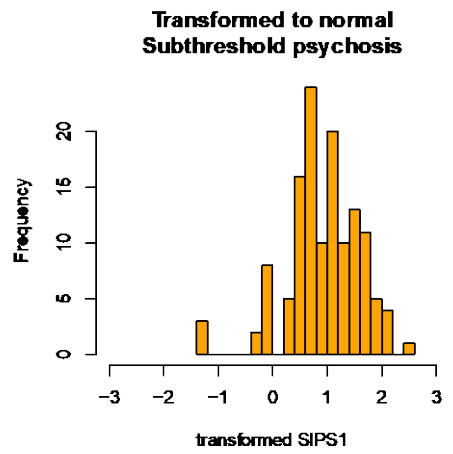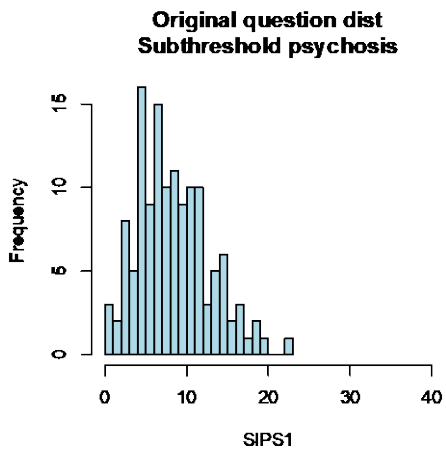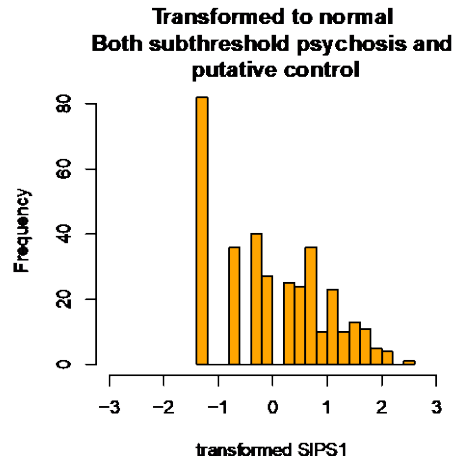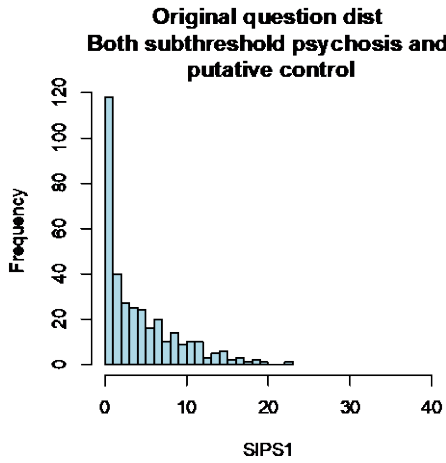
Supplementary Figure 7. Histogram of pre and post transformed SIPS measure. Transformation is defined by "qnorm(pexp(q = x + 0.5, rate = 0.2238))".
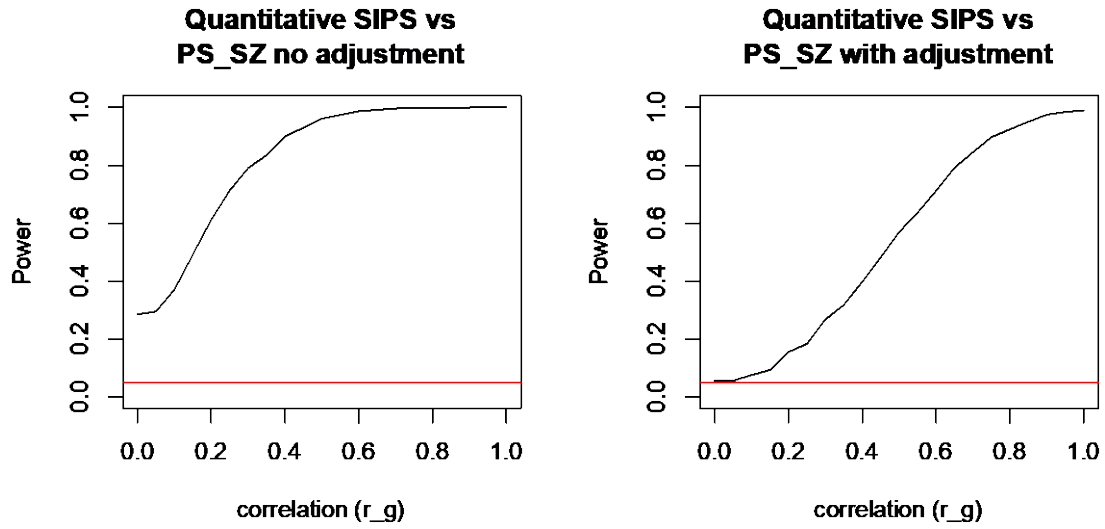
Supplementary Figure 8. Power analysis for quantitative subthreshold psychosis measure based on SIPS either without an adjustment for binary subthreshold psychosis (left) or with (right). Results were generated using simulation including only those simulations where a significant (i.e. alpha < 0.05) observation was made between subthreshold psychosis and PS_SZ. Note that the plot on the right, with the binary conditioning, is unbiased, unlike the plot on the left.
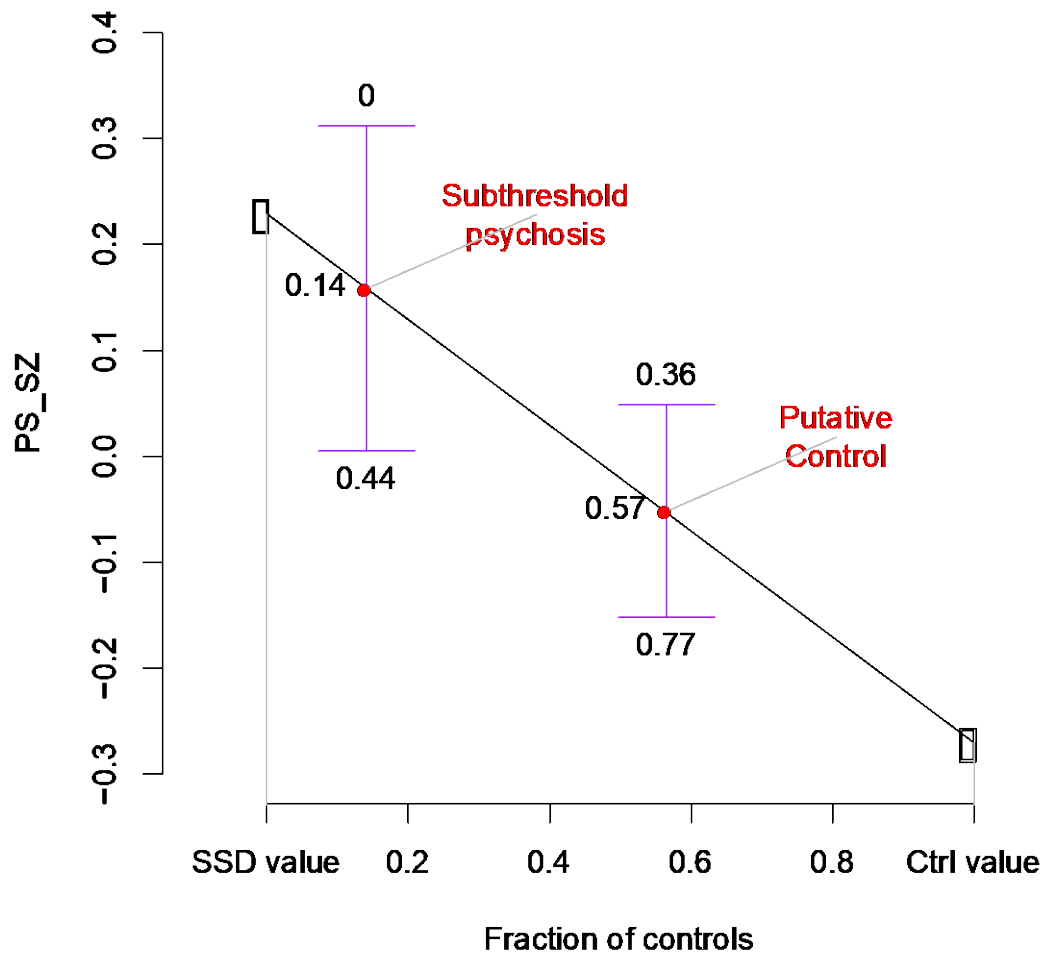
Supplementary Figure 9. In the subthreshold psychosis and putative control group we modelled the relative contribution of future SSD cases and true controls in each group, based on the observed PS_SZ (see Methods). The observed PS_SZ in the subthreshold group is consistent with a scenario in which 84% (95% CI 54 - 100%) of individuals who had subthreshold psychotic symptoms at the time of the assessment for this study would subsequently transition to SSD, a proportion inconsistent with known rates of SSD in 22q11DS.
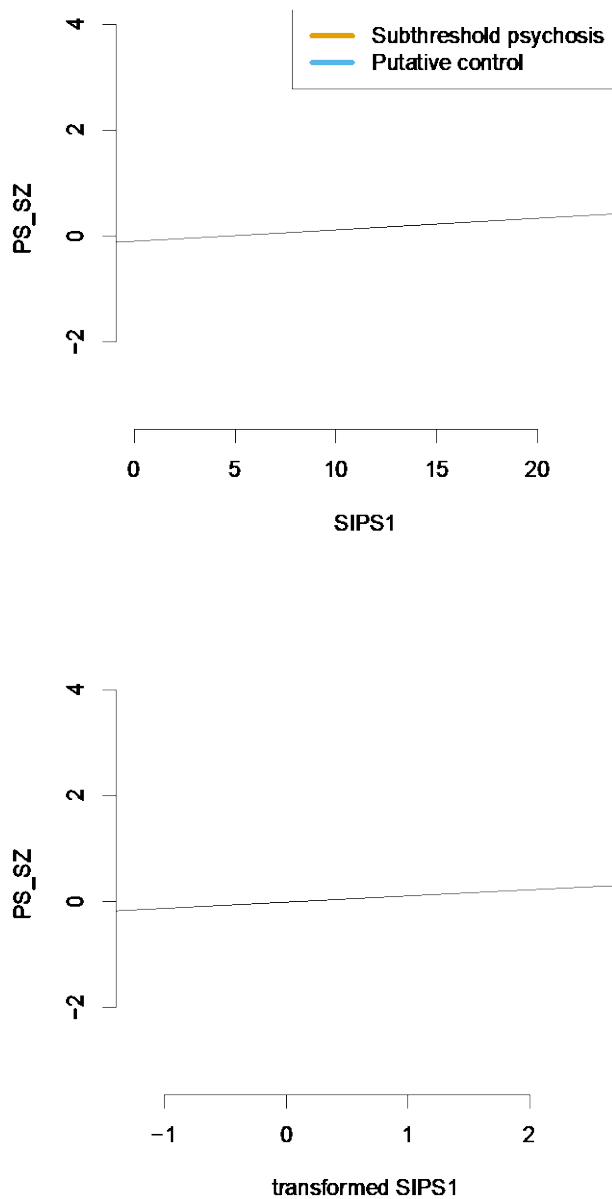
Supplementary Figure 10. Correlation plots between PS_SZ and a quantitative measure of subthreshold psychotic symptom severity. Upper panel shows untransformed SIPS values, lower panel shows transformed SIPS values. When adjusting for the previous binary indicator of subthreshold psychosis versus control, the association between the transformed quantitative SIPS phenotype and PS_SZ was not significant (N = 347, p = 0.77, r^2 = 0.0001).
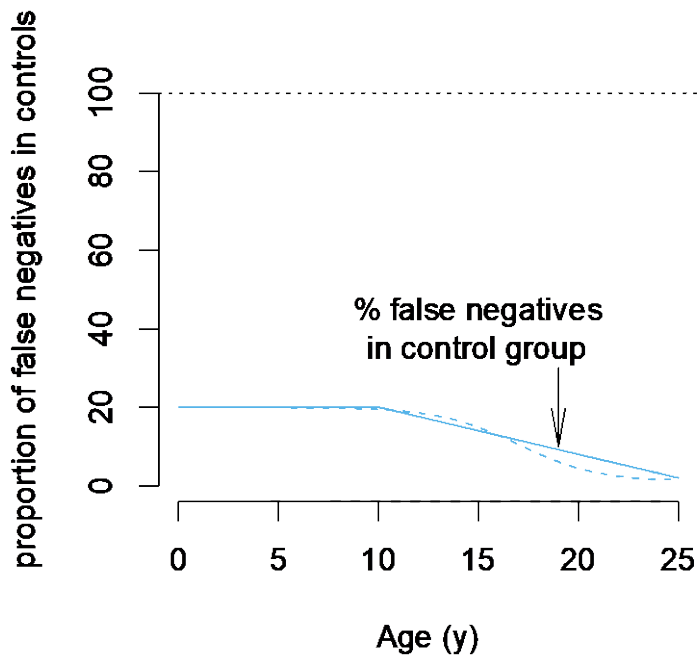
Supplementary Figure 11. Proportion of false negative controls over age range (0-25 years) (blue line). False negatives are subjects that will eventually develop schizophrenia but have not been identified as such at the time of this study. We assumed that the proportion of false negatives in the control group is equal to 20% at birth and declines as an inverse function of the cumulative incidence of schizophrenia. Although the decline follows a pseudo sigmoid function (dotted line) it can be approximated by a linear relationship.
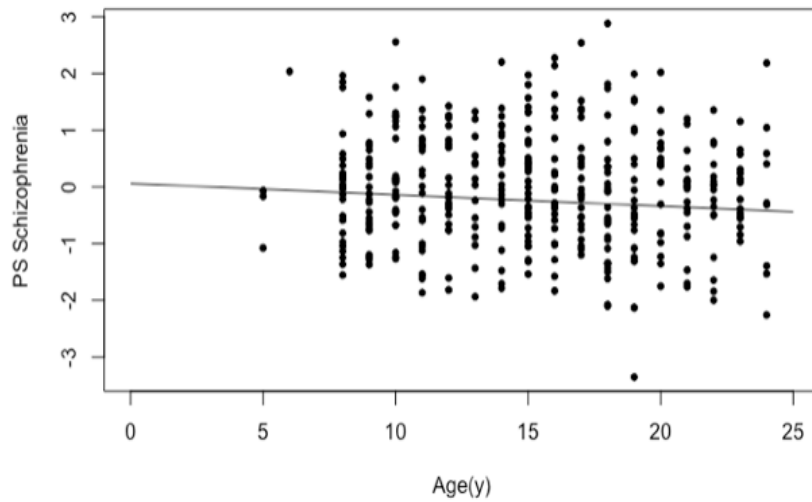
Supplementary Figure 12. Age shows a moderate linear correlation with PS_SZ in the putative controls group (r2 = -0.2, p =0.048).

1. Schneider, M.*, et al.* Ultra high risk status and transition to psychosis in 22q11.2 deletion syndrome. *World psychiatry : official journal of the World Psychiatric Association* **15**, 259-265 (2016).
2. Cross-Disorder Group of the Psychiatric Genomics, C.*, et al.* Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nature genetics* **45**, 984-994 (2013).
3. Sieradzka, D.*, et al.* Are genetic risk factors for psychosis also associated with dimension-specific psychotic experiences in adolescence? *PloS one* **9**, e94398 (2014).
4. Jones, H.J.*, et al.* Phenotypic Manifestation of Genetic Risk for Schizophrenia During Adolescence in the General Population. *JAMA psychiatry* **73**, 221-228 (2016).
5. Jones, H.J.*, et al.* Investigating the genetic architecture of general and specific psychopathology in adolescence. *Transl Psychiatry* **8**, 145 (2018).

Abstract for 2020 World Congress of Psychiatric Genetics

Title (200 characters max excl spaces > 98): Using common genetic variation to examine phenotypic expression and risk prediction in 22q11.2 Deletion Syndrome

Abstract (3000 characters max excl spaces > 2411):
Background: The onset of schizophrenia is often preceded by a decline in IQ, and subthreshold symptoms of psychosis, and low intellectual functioning (IQ) early in life is also a risk factor. However, it remains currently unclear whether these schizophrenia-related phenotypes share a substantial genetic basis with schizophrenia, or intellectual functioning. Our main objectives were twofold. First, to study the genetic relationship between schizophrenia and schizophrenia-related phenotypes (i.e., low baseline IQ, cognitive decline and subthreshold positive psychotic symptoms). Second, to examine the use of polygenic scores for schizophrenia and IQ for individual risk prediction of schizophrenia and intellectual disability in a population with an *a priori* increased risk for these phenotypes: individuals with 22q11.2 Deletion Syndrome (22q11DS; associated with a 20-25% risk for schizophrenia and a 40-50% risk for intellectual disability).

Methods: Using data from the International 22q11.2 Deletion Syndrome Brain Behavior Consortium (IBBC), we examined phenotypic and genetic data in a cohort of 962 individuals with 22q11DS. We used Wechsler scales to assess baseline IQ and (using longitudinal data) IQ-decline, and standardized clinical measures of psychosis to assess subthreshold psychotic status. First, we examined the association of these phenotypes with two polygenic scores, respectively for schizophrenia and intelligence, derived from the general population. Second, we evaluated their use for individual risk prediction in 22q11DS.

Results: The polygenic scores for schizophrenia and intelligence were not only associated with schizophrenia and baseline IQ respectively, but the schizophrenia polygenic score was also significantly associated with cognitive (verbal IQ) decline and nominally associated with subthreshold psychosis. Further, comparing the tail-end deciles of the schizophrenia and IQ polygenic score distributions, 33% versus 9% of 22q11DS subjects had schizophrenia, and 63% versus 24% had intellectual disability.

Discussion: First, the results demonstrate shared genetic underpinnings of schizophrenia-related phenotypes with schizophrenia, suggesting that cognitive decline and subthreshold psychosis are not merely phenotypically associated with schizophrenia, but may represent early stages of the illness. Second, our findings demonstrate that in 22q11DS, polygenic scores enable stratification in high and low risk groups substantially in excess of what would be found in the general population, highlighting the potential of polygenic scores for risk stratification among individuals with highly, but incompletely, penetrant variants.

Authors:
Robert W. Davies[1,2*], Ph.D, Ania M. Fiksinski[3,4*], M.Sc., Elemi J. Breetvelt[5], M.D., Ph.D., Nigel M. Williams[6], Ph.D., Stephen R. Hooper[7], Ph.D., Thomas Monfeuga[6], Ph.D., Anne S. Bassett[4,8,9], M.D., FRCPC, Michael J. Owen[6], M.D., Ph.D., Raquel E. Gur[10], M.D., Ph.D., Bernice E. Morrow[11], Ph.D., Donna M. McDonald-McGinn[12,13], M.Sc., LCGC., Ann Swillen[14,15], Ph.D., Eva W. C. Chow[4,8], M.D., Marianne van den Bree[6], Ph.D., Beverly S. Emanuel[12], Ph.D., Joris R Vermeesch[14], Ph.D., Therese van Amelsvoort[16], M.D., Ph.D., Celso Arango[17], M.D., Ph.D., Marco Armando[18], M.D., Ph.D., Linda E.

Campbell[19], Ph.D., Joseph F. Cubells[20,21], M.D., Ph.D., Stephan Eliez[18], M.D., Sixto Garcia-Minaur[22], M.D., Doron Gothelf[23,24], M.D., Wendy R. Kates[25], Ph.D., Kieran C. Murphy[26], M.D., Ph.D., Clodagh M. Murphy[27], MD, Ph.D., Declan G. Murphy[27], M.D, FRCPsych, Nicole Philip[28,29], M.D., Gabriela M. Repetto[30], M.D., Vandana Shashi[31], MBBS., M.D., Tony J Simon[32], Ph.D., Damiàn H. Suñer[33], Ph.D., Stefano Vicari[34], M.D., Stephen W. Scherer[36], Ph.D., International 22q11.2 Brain and Behavior Consortium[#], Carrie E. Bearden[35], Ph.D., Jacob A.S. Vorstman[3,36,37], M.D., Ph.D.

* These authors contributed equally to this work
# Group authorship; authors to be listed in the acknowledgement section at the end of this paper

Affiliations:
1. Program in Genetics and Genome Biology and The Centre for Applied Genomics, The Hospital for Sick Children, Toronto, Ontario, Canada.
2. Department of Statistics, University of Oxford, 24-29 Giles' St., Oxford, OX1 3LB, United Kingdom
3. Department of Psychiatry, Brain Center, University Medical Center Utrecht, the Netherlands.
4. The Dalglish Family 22q Clinic, Clinical Genetics Research Program, Centre for Addiction and Mental Health, Toronto, Canada.
5. Department of Psychiatry, The Hospital for Sick Children, Toronto, Ontario, Canada
6. MRC Centre for Neuropsychiatric Genetics and Genomics, Division of Psychological Medicine and Clinical Neurosciences, School of Medicine, Cardiff University, Cardiff, United Kingdom.
7. Department of Allied Health Sciences, School of Medicine, University of North Carolina-Chapel Hill, North Carolina, United States of America.
8. Department of Psychiatry, University of Toronto, Ontario, Canada.
9. Toronto General Hospital, University Health Network, Toronto, Ontario, Canada.
10. Department of Psychiatry and Lifespan Brain Institute, Penn Medicine - CHOP, University of Pennsylvania, Pennsylvania, United States of America.
11. Department of Genetics, Albert Einstein College of Medicine, Bronx, New York, United States of America.
12. Division of Human Genetics and 22q and You Center, the Children's Hospital of Philadelphia, Pennsylvania, United States of America.
13. Department of Pediatrics, Perelman School of Medicine of the University of Pennsylvania, Philadelphia, United States of America.
14. Center for Human Genetics, University of Leuven (KU Leuven), Belgium.
15. Department of Human Genetics KU Leuven, Belgium.
16. School Mental Health & Neuroscience Maastricht University, the Netherlands.
17. Department of Child and Adolescent Psychiatry, Hospital General Universitario Gregorio Marañón, IiSGM, CIBERSAM, School of Medicine, Universidad Complutense, Madrid, Spain
18. Developmental Imaging and Psychopathology, Department of Psychiatry, University of Geneva, Switzerland.
19. School of Psychology, University of Newcastle, Newcastle, Australia.
20. Department of Human Genetics, Emory University School of Medicine, Atlanta, Georgia, United States of America.
21. Emory Autism Center, Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, Atlanta, Georgia, United States of America.

22. Institute of Medical and Molecular Genetics (INGEMM), La Paz University Hospital, Madrid, Spain.
23. The Child Psychiatry Division, Edmond and Lily Safra Children's Hospital, Sheba Medical Center, Tel Hashomer, Israel.
24. Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel.
25. Department of Psychiatry and Behavioral Sciences, SUNY Upstate Medical University, New York, United States of America.
26. Department of Psychiatry, Royal College of Surgeons in Ireland, Beaumont Hospital, Dublin, Ireland
27. Department of Forensic and Neurodevelopmental Sciences, Institute of Psychiatry, Psychology & Neuroscience (IoPPN), King's College London, United Kingdom.
28. Département de Génétique Médicale, APHM, CHU Timone Enfants, Marseille, France.
29. Aix Marseille Université, MMG, INSERM, Marseille, France.
30. Centro de Genética y Genómica, Facultad de Medicina, Clínica Alemana Universidad del Desarrollo, Santiago, Chile.
31. Department of Pediatrics, Division of Medical Genetics, Duke University School of Medicine, North Carolina, United States of America.
32. MIND Institute and Department of Psychiatry and Behavioral Sciences, University of California Davis, Sacramento, California, United States of America.
33. Laboratorio Unidad de Diagnóstico Molecular y Genética Clínica, Hospital Universitari Son Espases, Palma de Mallorca, Spain.
34. Child and Adolescence Neuropsychiatry Unit, Department of Neuroscience, IRCSS Bambino Gesù Children's Hospital of Rome, Italy.
35. Departments of Psychiatry and Biobehavioral Sciences and Psychology, Semel Institute for Neuroscience and Human Behavior. University of California, Los Angeles, California, United States of America.
36. Program in Genetics and Genome Biology, SickKids Research Institute, Toronto, Ontario, Canada.
37. Department of Psychiatry, The Hospital for Sick Children, University of Toronto, Ontario, Canada.