Neuropsychiatric risk in children with intellectual disability of genetic origin: IMAGINE - The UK National Cohort Study

Jeanne Wolstencroft, PhD[1]; Francesca Wicks, BSc[2,8]; Ramya Srinivasan, MD [1,9]; Sarah Wynn[8]; Tamsin Ford, PhD[2]; Kate Baker, PhD[2,3,10] Samuel J R A Chawner, PhD[5,6,7], Jeremy Hall, MD[5,6]; Marianne B M van den Bree, PhD[5,6]; Michael J. Owen, PhD[5,6]; IMAGINE Study[*] ; David Skuse, MD [1]; F. Lucy Raymond MD [2,3,4].

Affiliations:

1 NIHR BRC Great Ormond Street Institute of Child Health, University College London, London, UK.

2 School of Clinical Medicine, University of Cambridge, Cambridge, UK

3 Cambridge University Hospitals NHS Foundation Trust, Cambridge Biomedical Campus, Cambridge, UK

4 NIHR Bioresource, Cambridge Biomedical Campus, Cambridge, UK

5 Medical Research Council Centre for Neuropsychiatric Genetics and Genomics, Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff, UK.

6 Neuroscience and Mental Health Research Institute, Cardiff University, Cardiff, UK.

7 Cardiff University Centre for Human Developmental Science, School of Psychology, Cardiff University, Cardiff, UK

8 UNIQUE Charity, Oxted, UK

9 UCL Division of Psychiatry, University College London, London, UK

10 MRC Cognition and Brain Sciences Unit, University of Cambridge, UK

* The members of the IMAGINE Study are listed in Appendix 4

Corresponding author

Prof F Lucy Raymond

flr24@cam.ac.uk

44 7824994699

School of Clinical Medicine, University of Cambridge, Cambridge, UK

Abstract/Summary

Background

Children with intellectual disability (ID) frequently have multiple co-morbid neuropsychiatric conditions and poor physical health. Genomic testing is increasingly recommended as a first-line investigation for these children. We aimed to determine the impact of genomics, inheritance and socioeconomic deprivation on neuropsychiatric risk in children with intellectual disability of genetic origin as compared to the general population.

Methods

IMAGINE is a prospective study using online mental health and medical assessments in a cohort of 2770 children with ID and pathogenic genomic variants, identified by the UK's National Health Service.

Outcomes

Assessments completed on 2397 young people with ID (4-19 years, M 9·2, SD 3·9) with a rare pathogenic genomic variant. 1339 (55·9%) were male and 1058 (44·1%) were female. 1771 (73·9%) of participants had a pathogenic copy number variant (CNV), 626 (26·1%) a pathogenic single nucleotide variant (SNV). Participants were representative of the socioeconomic spectrum of the UK general population. The relative risk of co-occurring neuropsychiatric diagnoses, compared with the UK national population, was high: Autism Spectrum Disorder 29·2 (95% CI 23·9 to 36·5), Attention Deficit Hyperactivity Disorder 13·5 (95% CI 11·1 to 16·3). In children with a CNV, those with a familial variant tended to live in more socioeconomically deprived areas. Both inheritance and socioeconomic deprivation contributed to neuropsychiatric risk in those with a CNV.

Interpretation

Children with genomic variants and ID are at a greatly enhanced risk of neuropsychiatric difficulties. CNV variant inheritance and socioeconomic deprivation also contribute to the risk. Early genomic investigations of children with intellectual disability could facilitate the identification of the most vulnerable children. In addition, harnessing parental expertise using online DAWBA assessments could rapidly highlight children with exceptional needs to child mental health services.

Funding

UK Medical Research Council and Medical Research Foundation.

Research in context

Evidence before the study

We searched PubMed title/abstract for publications in English from database inception until June 11, 2021, using the search terms ((child*) AND ((developmental delay) OR (intellectual disability)) AND (mental health) AND (cohort). Only one national cohort survey of children's mental health has reported on the increased risk of mental health and neurodevelopmental disorders among children with ID using standardized measures. Other relevant cohort studies have focused on the identification of specific neurodevelopmental disorders (e.g., ASD, ADHD) in a population, which may incidentally be associated with ID, but have not reported on co-occurring behavioural or emotional problems. One genotype-first study of developmental delay in non-syndromic children has been published, but this did not systematically evaluate neurodevelopmental risk or mental health.

Added value of this study

Our nationally representative cohort of children (4-19 years, M 9·2, SD 3·9; 55·9% male, 44·1% female) with identified pathogenic genomic variants encompassing CNVs and SNVs that are far more varied than any previous genotype-first investigation of neurodevelopmental risk. Data were collected using standardized measures of child mental health that are equivalent to those used in UK national surveys and thus allow direct comparison with general population data collected contemporaneously. Previous studies of neuropsychiatric risk in children with ID have either been small-scale cohorts or were not designed to evaluate such a wide range of mental health issues. Those that have sought evidence for genetic predisposition have, in almost all instances, started from a phenotype of interest (such as ASD) and then screened for pathogenic variants. The unique contribution of this investigation is that it provides evidence from a genotype-first investigation of neuropsychiatric risk, with the predisposing genomic variants reported by a UK NHS diagnostic protocol. The addition of data on socioeconomic status is based on a multifaceted UK index of multiple deprivation; a variable that has not previously been used in epidemiological studies of mental health risk in children with developmental delay (DD).

Implications of all the available evidence

Routine genomic testing is identifying pathogenic variants in an increasing proportion of

children with developmental delay but, with the exception of a few relatively well-studied variants, the implications of a genomic disorder for a child's future mental health is currently unknown. Intellectual disabilities are generally associated with an increased risk of neurodevelopmental disorders, as well as emotional and behavioural problems, but this study has shown that in children whose DD has an identifiable genetic aetiology that the risk is amplified considerably. Our findings have implications for the clinical management of such children and indicate an urgent need for early assessment and intervention

Introduction

The genomic basis of intellectual disability (ID) is being unveiled at pace. Large-scale identification of highly penetrant variants that cause developmental delay, ID and autism has been achieved using next generation sequencing methods with a trio-based design (parents and child) [1-3]. Best practice guidelines recommend exome or genome sequence analysis as a first or second tier investigation for all children presenting with developmental delay or ID [4]. Yet, with a few exceptions, the confidence with which a rare genomic variant can be regarded as pathogenic is not matched by an equivalent confidence about the implications of that finding for the child's future neuropsychiatric profile. Compared with the general population, children with ID have significant additional needs, in terms of physical and mental health, but we know little about the influence of most identified rare genomic variants on a child's long-term outcome[5]. Most previous studies of children with ID are limited by modest sample sizes[5,6] or have selected participants from large epidemiological cohorts in which the proportion of children with moderate to profound ID was small[7,8]. So far, no national cohort study of ID has collected genomic data at scale. We do not know to what extent prognosis is influenced by environmental factors such as socioeconomic deprivation or family factors, including the inheritance of the genomic variant (familial or *de novo*). This study was designed to assess social and demographic influences on the physical and mental health of a national cohort of children and young people with rare genomic disorders associated with ID. It was also designed to make a comparison of prevalence with the equivalent diagnostic data provided by the UK National Survey of Children's Mental Health 2017 (https://digital.nhs.uk/data-and-information/publications/statistical/mental-health-of-children-and-young-people-in-england/2017/2017). If physical and mental health care needs can be predicted at the point of genetic diagnosis, then early personalised interventions may benefit the most vulnerable children.

Methods

Study Design and participants

3407 participants were recruited to the IMAGINE study (Intellectual Disability and Mental Health: Assessing the Genomic Impact on Neurodevelopment, https://imagine-id.org/) between 1/10/2014 and 30/06/2019 under London Research Ethics Committee-Queen Square 14/LO/1069. The criteria for entry into the study were: 1) presence of developmental delay or ID diagnosis made by a clinical care team 2) a confirmed molecular genetic diagnosis documented from an accredited diagnostic laboratory and 3) age of at least 4 years at enrolment. Recruitment to the study was by referral from all UK Regional Genetics Centres (94·4%), self-referrals or patient support groups (5·6%). This study is focused on a subset of the sample who were between 4 and 19 years of age. A parent or guardian provided consent on behalf of children under 16 years. All children received a copy of a storybook *Avery* written for the this study to facilitate a parent-child discussion about involvement in research[9]. For individuals >16 years who lacked capacity, consultees acted on their behalf.

Procedures

Diagnostic genomic reports were obtained from NHS medical records or directly from families including inheritance information where available (Appendix pp.2). Pathogenic variants were classified according to American College of Medical Genetics (ACMG) guidelines and only those with pathogenic or likely pathogenic variants were included[10]. For individuals with multiple genetic variants, subsequent data analysis was based on the most pathogenic variant.

Choice of primary outcomes

Primary caregivers were invited to complete online assessments of their child's educational progress, physical and mental health. We chose the *Development and Well-Being Assessment (DAWBA)* and the *Strengths and Difficulties Questionnaire (SDQ)* as our primary outcome measures (Appendix pp.3)[11,12], as both have been used in national studies of children's mental health in the UK. The potentially lengthy comprehensive psychiatric interview (DAWBA) provides DSM-5 compatible diagnoses, and broader measures of adjustment and family functioning too. It has been used for both UK national [8] and international surveys [12] of mental health in children from 5-17 years (to 19 years in latest survey). Making a clear distinction between problem behaviour in general and specific psychiatric disorders, is important in the ID population [7,13,14]. To maximise validity and reliability, rating procedures identical to the latest UK national survey were used[11].

Diagnoses, using DSM-5 criteria, were assigned by two independent experienced clinicians (Appendix pp.3)[15]. Inter-rater reliability was checked by co-rating 147 randomly chosen participants with the team that conducted the UK National Survey of Children's Mental Health (https://digital.nhs.uk/data-and-information/publications/statistical/health-survey-for-england/2018), and all kappa values for diagnostic categories were >0·7.

The *Strengths and Difficulties Questionnaire (SDQ)* assessed children's emotional and behavioural adjustment in dimensional terms [8,13,16,17]. The SDQ has been validated with children with IDD[13]. The SDQ includes scales that measure emotional symptoms, conduct problems, hyperactivity/impulsivity and inattention difficulties, peer relationship problems and prosocial behaviour. The first four scales are combined to create a total difficulties score. High scores are indicative of greater levels of mental health difficulty and scores above the 90th percentile indicate a high probability of a diagnosable psychiatric disorder (≥17) [17].

Secondary outcomes

Daily living skills were measured using the *Adaptive Behaviour Assessment System 3 (ABAS-3)*[18]. A developmental quotient was calculated from parental estimates of the child's mental age divided by their chronological age[7,19]. General physical health was estimated using parent ratings on the DAWBA (5 point Likert scale from very bad to very good). A structured questionnaire gathered information about pregnancy, birth, early development and current medical problems and medication.

Postcodes of participating family homes were scored on an Index of Multiple Deprivation (IMD), provided by the UK Office for National Statistics (https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019). In the current English Indices of Deprivation 2019 (IoD2019) seven domains of deprivation are considered and weighted as follows; Income (22·5%), Employment (22·5%), Education (13·5%), Health (13·5%), Crime (9·3%), Barriers to Housing and Services (9·3%) and Living Environment (9·3%). The indices of multiple deprivation for Wales, Scotland, England, and Northern Ireland are calculated separately.

Data analysis

Four sets of analyses were conducted. First, descriptive statistics were computed to describe the cohort's characteristics in the following domains: genetics; development, education and adaptive impairment; socio-economic status; and neuropsychiatric risk. Secondly, we conducted group comparisons using chi squared tests on the prevalence of DAWBA diagnoses between the IMAGINE cohort and the UK national study. Then, we conducted the

third and fourth set of analyses on a subset of the cohort who had a CNV of known inheritance (i.e. *de novo* or familial status). The third analysis compared the behavioural phenotypes and neuropsychiatric risk of children based on the inheritance of their CNV (familial/de novo). The Bonferroni method was used to adjust the threshold of significance for multiple comparisons in the second and third set of analyses. Our fourth and final set of analyses investigated the association between variant. Inheritance (*de novo* or familial status), the indices of deprivation (IMD) and the severity of behavioural and emotional difficulties (SDQ) using multivariable hierarchical linear regressions. Model 1 predicted the degree of behaviour difficulties (SDQ total score) from the IMD quintile and variant inheritance (binary variable: *de novo*/familial). Model 2 adjusted for confounds including sex, age of diagnosis, developmental quotient (developmental age/chronological age) and physical health problems (rated by parents on a 5 point likert scale from very bad to very good). Model 3 added an interaction factor (deprivation x inheritance). All data were analysed in SPSS version 24[20].

Role of the funding source

The study funders and sponsors were not involved in the study design, the collection, analysis, and interpretation of data, in the writing of the report and in the decision to submit the paper for publication.

Results

Participants

A total of 3407 participants were recruited to the IMAGINE study[19,21] (Figure 1). 2770 were aged between 4 to 19 years. Of these, 2397 (86·5%) families completed basic assessments of their child's mental health (Figure 1; Appendix pp.4). The mean age in this subsample was 9·2 years (SD 3·9) and 1339 (55·9%) were male and 1058 (44·1) were female.

Genetics

The cohort represented a diverse group of 2770 individual children with many different genomic disorders, 2397 of whom have measures of mental health available (Figure 1). 1771 (73·9%) individuals had a pathogenic copy number variant (CNV) and 626 (26·1%) individuals had a pathogenic single nucleotide variant (SNV). Familial variants were identified in 645 (26·9%) individuals; *de novo* variants were identified in 940 (39·2%); and in 812 (33·9%) individuals the parental results were not available to the study, thus, familial or *de novo* status could not be determined for the pathogenic variant (Appendix pp.4).

The average age at diagnosis of a pathogenic CNV was 5·4 years (SD 3·7). In total, 961 different CNV loci were observed within the cohort (Appendix pp.4). Where the inheritance was known, 564 (51%) individuals with a CNV had a familial variant compared to 541 (49%) who had a *de novo* CNV.

The average age at diagnosis of a pathogenic SNV was 7·8 years (SD 4·2). Pathogenic variants in 205 different single genes were observed (Appendix pp.2). Where the inheritance was known, most SNVs 399 (83·1%) were documented as de novo, compared to 81 (16·9%) that were documented as familial .

Development, education and adaptive impairment

Most children in the cohort had delayed developmental milestones, according to parental reports; the average age at first walking unsupported was 23·2 months (SD 13·5) and 1735 (72·4%) had delayed language skills.

912 (38%) children attended specialised education units or schools, 953 (39·8%) attended mainstream school with classroom assistance, 165 (6·9%) attended mainstream school without allocated support, 111 (4·6%) were not at school and 256 (10·7%) the type of schooling was not documented. Supplemental information (1277; Appendix pp.3) indicated that 976 (76·4%) had Special Educational Needs or an Education Health Care Plan (https://www.gov.uk/government/statistics/education-health-and-care-plans-england-2021). 978 (76·6%) caregivers received a Disability Living Allowance for their child (https://www.gov.uk/disability-living-allowance-children).

The ABAS-3 was completed for 1238 children: 63 (5.1%) scored in the average range, 120 (9·7%) in the below average range, 238 (19·2%) in the low range and 817 (66%) in the extremely low range according to ABAS-3 norms[22].

Socio-economic status

Of the 2397 children for whom measures of mental health are available, residential postcodes linked to IMD scores were available on 2277 UK participants from 2123 households (https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/adhocs/13773populationsbyindexofmultipledeprivationimddecileenglandandwales2020). The distribution of IMD scores approximated a uniform distribution; the cohort was representative of the UK national population based on IMD quintiles (Table 1, Appendix pp.4). Households with children with a familial variant were significantly over-represented in socio- economically more deprived quintiles, whilst the opposite trend was observed in households with children with a *de novo* variant (Figure 2). Ethnicity data was not collected.

Neuropsychiatric risk

The Strength and Difficulties Questionnaire (SDQ) scores (2397) revealed a high prevalence of behaviour difficulties compared to the UK national survey norm (Table 1)[8]. Most of the sample (1992, 83·1%) scored above the clinical 'cutpoint', compared to 20% of the general population of equivalent age and sex. Of these children; 334 (13·9%) had slightly raised scores, 378 (15·8%) had high scores, and 1279 (53·4%) had very high scores. Subscale scores for specific traits were raised to an equivalent degree (Appendix pp.4).

Clinically significant neuropsychiatric disorders were observed in 1161 (53·1%) of the children with completed DAWBAs (2186), compared with 12·8% in the 2017 national population survey (https://digital.nhs.uk/data-and-information/publications/statistical/health-survey-for-england/2018; RR 4·1, 95%CI 3·9 to 4·5, p<·001). Autistic Spectrum Disorder (ASD) diagnostic criteria were met in 776 (35·5%) compared to 1·2% in the general population (RR 29·2, 95%CI 23·9 to 36·5, p<·001). Attention deficit hyperactivity disorder (ADHD) diagnostic criteria were met in 473 (21·6%) compared with 1·6% in the general population (RR 13·5, 95%CI 11·1 to 16·3, p<·001, Table 2). Oppositional defiant disorders (ODD) were also relatively common (264, 12·1% vs 222, 2·9%; RR 4·2, 95%CI 3·5 to 5, p<·001), but the rates of conduct disorder were not raised (34, 1·6% vs 130, 1·7%). Anxiety disorders were identified in 232 (10·6%) compared with 551 (7·2%) in the comparison cohort (RR 1·5, 95%CI 1·3 to 1·7, p<·001, Table 2). Rates of depression were significantly lower, 9 (0·4%) vs 161 (2·1%) (RR 0·2, 95%CI 0·1 to 0·4, Table 2). Of the 1161 (53·1%) who met criteria for any psychiatric diagnosis, 483 (41·6%) had two or more co-occurring disorders, of which the most frequent co-occurring conditions were ASD and ADHD (247, 21·3%).

Physical health disorders

1277 (46·1%) caregivers completed a supplemental medical history questionnaire. 1195 (93·6%) reported at least one significant physical health problem. Many children had a history of seizures (355, 29·7%), the most common were absence seizures (148/355, 41·7%), generalized tonic-clonic seizures (120/355, 33·8%) and febrile seizures (94/355, 26·5%, (Appendix pp.5). Of those with a history of seizures, 188/355 (53%) were on specific anti-epileptic medication(s). Other physical health problems were common: 825 (64·6%) reported disturbed sleep; motor or movement disorders affected 814 (63·7%) of participants; 587 (46%) had fine motor control problems; 24 (1·9%) had cerebral palsy (Appendix pp.5).

Neuropsychiatric risk in children with CNVs by variant inheritance (*de novo* /familial)

Variant inheritance was examined for its contribution to risk of neuropsychiatric disorder in the CNV group (*de novo* 541, familial 564). Too few familial SNVs were observed for comparison (*de novo* 399, familial 81) (Appendix pp.3). Children with a *de novo* CNV variant were more impaired in their intellectual function, but not in their adaptive functioning compared to those with a familial variant, (DQ t(784·7)=4·4, p<·001, 95%CI 0·04 to 0·1; ABAS (t(547)=1·9, p=·06, 95%CI -0·7 to 4·5; Table 3). In contrast, more severe behavioural and emotional problems were observed in participants with a familial variant (SDQ t(1103)=10.6, p<·001, 95%CI 3·4 to 4·9). Those with a familial (p<·001, 95%CI 3·4 to 4·9). Those with a familial variant were also at a higher risk of specific mental health diagnoses, including ASD and ADHD, than those with a *de novo* variant (ASD RR 1·6 95%CI 1·4 to 1·9. p<·001; ADHD RR 1·9 95%CI 1·5 to 2·5 p<·001) and they were more likely to live in more deprived socio-economic areas (Table 3).

Genomic and Socioeconomic contributions to neuropsychiatric conditions

In children with a CNV, there was a greater severity of behavioural and emotional disorders, in those whose variant was familial and there was an association with socio-economic deprivation. In addition, there was a significant shift in the sex ratios of those with familial variants (Figure 3). The cohort overall was 1339 (55·9%) male and 1058 (44·1%) female whilst those with familial genomic variants were 356 (63·0%) male and 209 (37·0) female.

We conducted a series of hierarchical multivariable linear regressions to test the statistical significance of this association. Model 1 predicted the degree of behaviour difficulties (SDQ total score) from the IMD rank and variant inheritance (*de novo*/familial). Greater socioeconomic deprivation and possession of a familial variant both contributed to behaviour difficulties ($F(2, 803) = 56·7$, $p < ·001$, $b_{IMD=}$ -0·48, SE=0·16, p=·003; $b_{inheritance=}$ 4, SE=0·46, p<·0001). Model 2 adjusted for confounders including sex, age of diagnosis, developmental quotient and physical health ($F(6, 799) = 24·5$, $p < ·001$, Table 4); inheritance and the degree of deprivation remained predictors of behaviour difficulties ($b_{IMD=}$ -0·34, SE=0·16, p=·033; $b_{inheritance=}$ 3·7, SE=1·16, p<·001). Model 3 added an interaction factor (deprivation x inheritance). No significant interaction was found between the index of multiple deprivation and inheritance of the genomic variant (p=·41;Table 4).

Discussion

Our study, which involved over 2500 children, highlighted that intellectual disability of identifiable genetic aetiology is strongly associated with neurodevelopmental and mental

health disorders, and that the risk is greater in those whose genetic disorder is inherited, even after adjusting for developmental level, sex and socio-economic deprivation.

Our unique approach to measurement allowed us to include the assessment of disorders that are typically not included in studies of genetic risk in childhood. Former studies have either focused almost entirely on the physical consequences of genetic changes [2], or they have taken a relatively homogeneous population with a specific neurodevelopmental disorder (such as autism) and sought evidence of specific genomic variants that could have played a causal role[23]. Whilst we found that neurodevelopmental disorders were particularly frequently associated with ID of genetic origin, we also discovered that anxiety and oppositional defiant behaviour were major concerns. Previous studies that have examined the impact of pathogenic CNVs on child mental health have been small scale, focused on specific neurodevelopmental disorders (such as autism or schizophrenia), and considered only a small range of genomic variants. The IMAGINE study comprised a far wider range of CNVs, and a greater breadth of neuropsychiatric phenotypes, than any previous investigation of its type.

Consistent with previous work on intellectually disability in populations of children [7], we found an association between the degree of children's emotional and behavioural disturbance and families living in greater socioeconomic deprivation (https://www.instituteofhealthequity.org/resources-reports/a-fair-supportive-society-summary-report). Our first novel discovery was that such disorders were more prevalent among children whose genetic condition was inherited. The measurable impact of heritable variants on associated risk was largely confined to CNVs because SNVs were usually *de novo* in origin. Individuals with SNVs were also disproportionately drawn from less socially disadvantaged families, unlike CNVs which were identified in a socioeconomically representative cohort.

Considering the important finding that children with an inherited CNVs are at far greater risk of neuropsychiatric disorders, it is feasible that some parents also may have a degree of cognitive impairment themselves, associated with their carrier status, and thus are at social and educational disadvantage [24]. This could explain the observation that such families live in conditions of greater multiple deprivation and would contribute to the association with non-specific emotional and behavioural problems [25]. But we also found that the neurodevelopmental disorders, ADHD and ASD, were nearly twice as prevalent among children whose CNV was inherited. This difference could reflect some factors that

influenced the pathogenicity of the associated CNV, and/or polygenic risk that was also inherited[26], or unmeasured environmental factors that the study did not capture. Consistent with previously published data we found a relative paucity of females with familial variants compared to males supporting the theory of neuroprotective effect of the female sex[27]. We found that children with ID of genetic aetiology are not only at high risk of mental health and neurodevelopmental disorders, but also ~30% had a seizure disorders and other complex physical health needs. The children with seizures or absences were not confined to those with genomic variants within known epilepsy genes or genomic loci but was associated with a wider range of genomic disorders than anticipated suggesting the presence of a seizure disorder was a more generalised phenomenon in children with ID[28].

Our study has some limitations. Recruitment was almost exclusively based on referrals initiated by UK NHS Regional Genetics Centres (RGCs). Families with a child in whom a pathogenic variant had been diagnosed were approached with information about the IMAGINE Study by RGCs, therefore the number of families that declined to take part is unknown. Initial genetic investigations in most children were due to developmental delay. Genetic testing due to suspected autism cannot be excluded, although autism alone is not an indication for genetic investigations under current NHS guidelines. All participants in the UK came through NHS testing routes and a diverse range of technologies was used to make genomic diagnoses. The high number of children with CNV reflects historic limitations in diagnostic technologies. The inheritance of each variant was only identifiable in 64% of participants.

The study did not include children with ID without a molecular diagnosis. It is unlikely that these children will have significantly different mental health needs compared to those with a genetic diagnosis, but our study could not inform this assumption.

Assessments of mental health were mostly obtained online and were based on parental report. It is possible that parents who have a rare genetic disorder themselves, are living in socioeconomically disadvantaged circumstances, provided a different account of their child's behaviour than those with no underlying rare genetic disorder who live in less deprived circumstances. To mitigate against parental bias in reporting, for example, cognitive levels, multiple validated and independent assessment tools were used throughout. In addition, the threshold for referral and difficulties navigating access to services including genetic testing may be far higher for children with a familial CNV. This would bias the sample to more severe neurodevelopment in children with a familial CNVs. Participation rates in those families who volunteered to join the cohort was very high: 85% completed at

least some assessments. A strength of our design was that we measured and assessed mental health and neurodevelopmental disorders using the same instruments used in other UK national studies, allowing direct comparisons with general population data. Our diagnostic evaluations were shown to be consistent with the diagnostic decision-making of the latest national UK survey of children's mental health. A further strength was that participants were recruited from the NHS genetic service which is free at the point of delivery and thus demographically and socioeconomically unbiased and provided consistent quality of diagnoses, based on accredited diagnostic reports.

Future research should evaluate the emergence of new mental health outcomes over time and investigate sex-differences in these trajectories. The median age of participating children was 9 years at our initial assessment, but serious mental health disorders that are associated with many of the genetic disorders we surveyed do not appear until adolescence or early adult life. We are now following up the families 5 years after our initial evaluation, in order to understand the impact of their genetic disorder on specific educational needs, and plan for appropriate medical management. At the point of a genetic diagnosis, often in very early childhood, this information is lacking for the majority of the conditions we identified. To conclude, we have conducted the largest survey yet of rare genomic variants that are associated with intellectual disability. The identification of a pathogenic CNV or SNV, in a child with developmental delay, indicates an exceptionally high risk of their developing an associated neurodevelopmental disorder or other mental health condition, irrespective of the specific rare genomic variant. Those in whom a genomic variant is inherited are particularly vulnerable. This information should be used to plan targeted assessments and interventions to support families at the earliest opportunity. Based on our data we would recommend better training for health care providers about the wider use and utility of genetic testing and its value in terms of predicting potential mental health needs of children. We would also recommend better use of parental expertise in pre-assessment of children's needs. Wider use of online assessments of children e.g. DAWBA could have a significant impact on identifying rapidly those children in most need of child mental health services which are currently hugely limited in the UK.

Data Sharing statement

The full phenotypic IMAGINE dataset is available from the UK Data Archive under special license access (SN 8621):

https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=8621

Requests for genotype or linked genotypic-phenotypic data can be made through the study's data access committee: https://imagine-id.org/healthcare-professionals/datasharing/

Declaration of Interests

No conflict of interest is disclosed by co-authors.

Author contribution statement

All authors contributed to writing the manuscript. JW conducted the analysis. JW, RS, FW, TF, FLR and DS verified data and had access to raw data. FLR had final responsibility for the decision to submit for publication

Acknowledgements

We wish to thank all the families and children who took part in the study. We also wish to thank all UK NHS Regional Genetics Centres and UNIQUE charity (https://rarechromo.org) for their time and support in offering referral and recruitment to the study to families. The work was supported by a grant from the MRC and MRF (grant no. MR/L011166/1 and MR/N022572/1). This research was supported by the NIHR and FLR and ERM were supported by Cambridge NIHR BRC. SJRAC is supported by an ISSF Wellcome Trust Fellowship (204824/Z/16/Z) and RS from the Wellcome Trust grant (211163/Z/18/Z). The University of Cambridge has received salary support in respect of EM and FLR from the NHS in the East of England through the Clinical Academic Reserve. The views expressed are those of the authors and not necessarily those of the NHS or Department of Health.

Figures

Figure 1: Cohort ascertainment flow chart.

The main recruitment source was UK Regional Genetic Centres 94·4% and the remaining 5·6% were ascertained through self-referral.

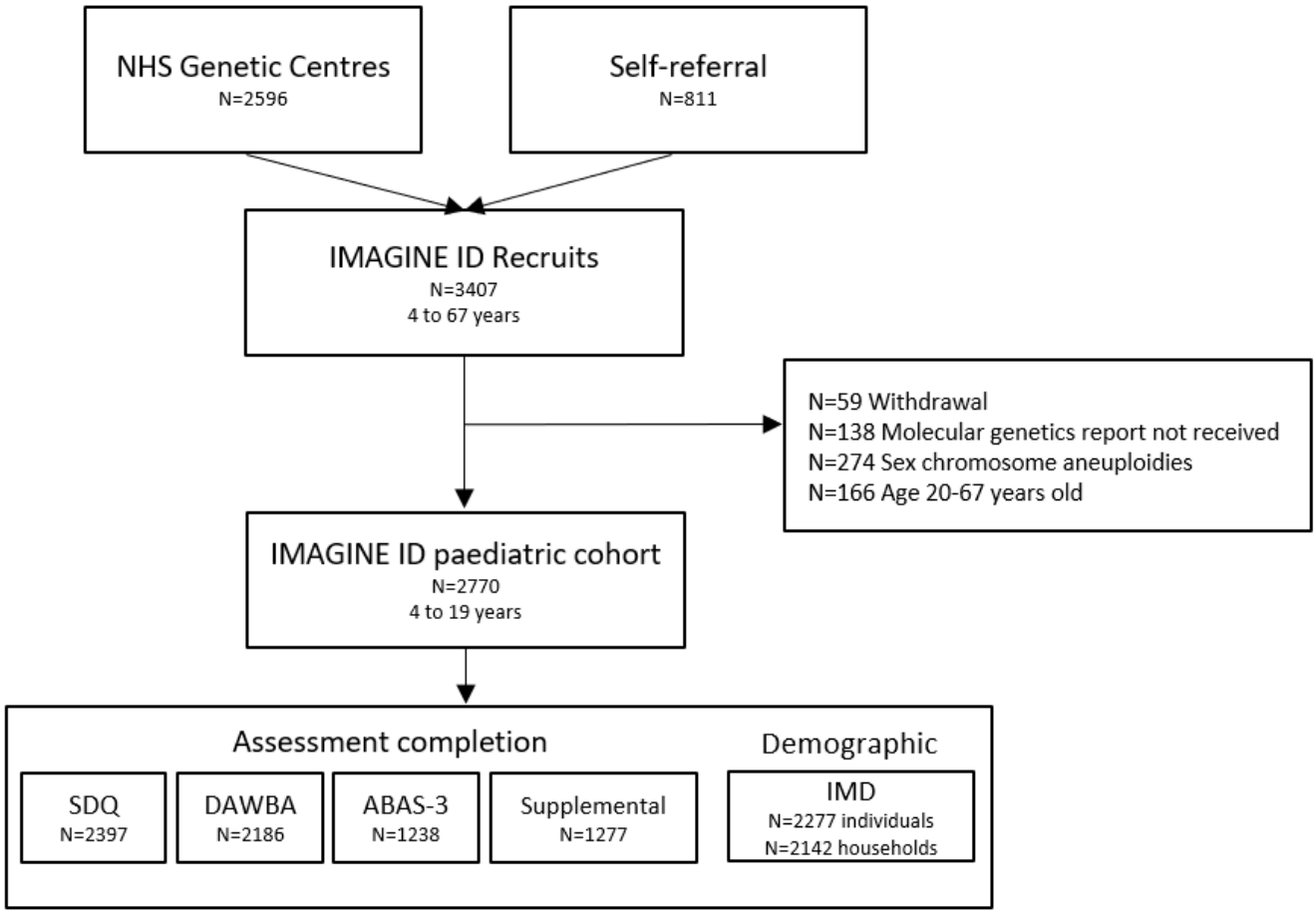


Figure 2: Inheritance by Index of Multiple Deprivation quintile

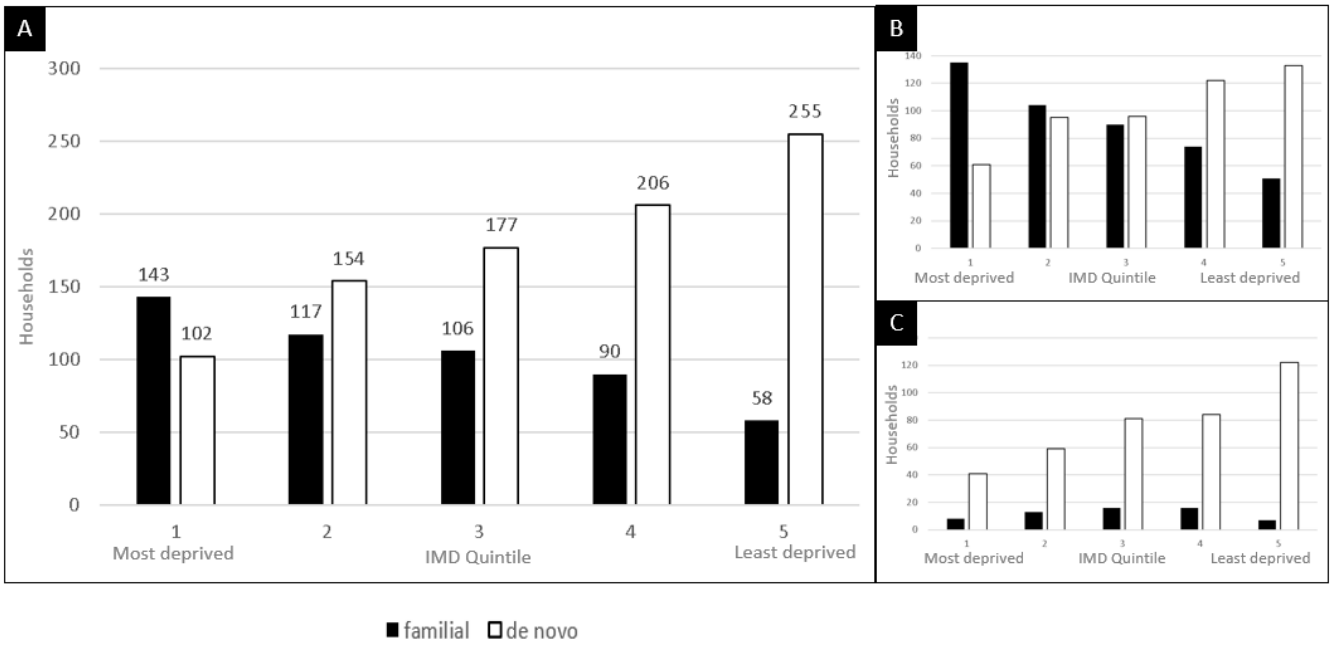
Panel A: Household IMD by inheritance for all variants of known inheritance (n=1408; $n_{familial}$=514, n_{denovo}=894; n_{male}=776, n_{female}=632)

Panel B: Household IMD by CNV variant inheritance (n=961; $n_{familial}$=454, n_{denovo}=507;n_{male}=546, n_{female}=415)

Panel C: Household IMD by SNV variant inheritance (n=447; $n_{familial}$=60, n_{denovo}=387; n_{male}=230, n_{female}=217)

IMD ranks by UK nations were combined to examine group differences between those households with an inherited and de novo variant. Households were scored once regardless of number of individuals within the household who had genetic variants. IMDs for variants of unknown significance are not represented ($n_{unknown}$=734). The 1st quintile includes the most deprived postcodes and the 5th quintile the least deprived postcodes.

Black bars = households with a familial variant; Open bars= households with *a de novo* variant.



Tables

Table 1: Cohort characteristics n=2397

Data are n (%) or specified. [1]n=2349; [2]n=2142; [3]n=1911, [4]n=2297; [5]n=1238. The developmental quotient was calculated from parental estimates of the child's mental age divided by their chronological age. General physical health was estimated using parent ratings on the DAWBA (5 point Likert scale from very bad to very good). See Appendix p4 for summary of n numbers.

Domain		No. (%)
Variant type	CNV	1771 (73·9)
	SNV	626 (26·1)
Variant inheritance	familial	645 (26·9)
	de novo	940 (39·2)
	not determined	812 (33·9)
Sex	Male	1339 (55·9)
	Female	1058 (44·1)
Age	4 to 8	1211 (50·5)

	9 to 11	531 (22·1)
	12 to 16	533 (22·2)
	17 to 19	122 (5·1)
Age at diagnosis[1]	Under 4	710 (30·2)
	4 to 8	1045 (44·5)
	9 to 11	341 (14·5)
	12 to 16	227 (9·7)
	17 to 19	26 (1·1)
IMD quintile by household[2]	1st	431 (20.1)
	2nd	406 (19.0)
1st - most deprived	3rd	407 (19·0)
5th -least deprived	4th	427 (19·9)
	5th	471 (22·0)
Intellectual functioning[3] **, mean (SD)**	Developmental quotient	0·55 (0·24)
	Mental age	4·94 (3)
Physical health[4]	Very good	635 (27·6)
	Good	944 (41·1)
	Fair	567 (24·7)
	Bad	119 (5·2)
	Very bad	32 (1·4)
ABAS-3 [5]	Extremely Low	817 (66)
	Low	238 (19·2)
	Below Average	120 (9·7)
	Average	62 (5)
	Above average	1 (0·1)
SDQ total score	Close to average	406 (16·9)
	Slightly raised	334 (13·9)
	High	378 (15·8)
	Very high	1279 (53·4)

Table 2: Neurodevelopmental and mental health diagnoses (n=2186)

IMAGINE N female = 960, N male= 1226

National study N female=3803, N male=3851

Data are n (%). IMAGINE DAWBA diagnoses were compared to the DAWBA prevalence estimates from the UK 2017 National study of child and adolescent mental health. Threshold of significance corrected for multiple comparisons using the Bonferroni correction method. P =·006

RR= Relative Risk; [1] Age 5-19 years; [2] chi^2 test of independence

DAWBA Diagnoses, N(%)	**IMAGINE n=2186**	**National study 2017** [1] **n=7654**	**P**[2]	**RR (95% CI)**

Emotional disorders Anxiety Depression	236 (10·8) 232 (10·6) 9 (0·4)	620 (8·1) 551 (7·2) 161 (2·1)	<·001 <·001 <·001	1·3 (1·2 to 1·5) 1·5 (1·3 to 1·7) 0·2 (0·1 to 0·4)
Behavioural disorders Oppositional defiant disorder Conduct disorder	283 (12·9) 264 (12·1) 34 (1·6)	352 (4·6) 222 (2·9) 130 (1·7)	<·001 <·001 ·71	2.8 (2·4 to 3·3) 4.2 (3·5 to 5) 0·9 (0·6 to 1·3)
Attention deficit hyperactivity disorder (ADHD)	473 (21·6)	123 (1·6)	<·001	13·5 (11·1 to 16·3)
Autism spectrum disorder (ASD)	776 (35·5)	92 (1·2)	<·001	29·2 (23·9 to 36·5)

Table 3: CNV group participant characteristic comparison by variant inheritance

Data are n (%) or as specified. [1]n=1106; [2]n=1098; [3]n=961; [4]n=1071, [5]n=855; [6]n=549;[7]n=1021;[8] DAWBA skip rules affect number of responses; [9] chi^2 test of independence; [10] Two-sample Kolmogorov–Smirnov test; Threshold of significance corrected for multiple comparisons using the Bonferroni correction method α=·002. The developmental quotient was calculated from parental estimates of the child's mental age divided by their chronological age (0= low developmental level, 1= high developmental level). General physical health was estimated using parent ratings on the DAWBA (5 point Likert scale from very bad to very good). IMD quintile reported by household.

See Appendix pp.4 for summary of n numbers.

	familial	**de novo**	**p**
Age[1], mean (SD)	8·7 (3·6)	8.9 (3·9)	·4
Age of diagnosis[2], mean (SD)	6 (3·6)	4.7 (3·9)	<·001
Sex (male) [1, 9]	356 (63.0)	281 (51.9)	<·001
Sex (female) [1, 9]	209 (37.0)	260 (48.1)	<·001
IMD quintile by household [3, 10] **1st - most deprived** **5th -least deprived**	1st 135 (29·7)	1st 61 (12)	<·001
	2nd 104 (22·9)	2nd 95 (18·7)	
	3rd 90 (19·8)	3rd 96 (18·9)	
	4th 74 (16·3)	4th 122 (24·1)	
	5th 51 (11·2)	5th 133 (26)	

Physical health[4, 8, 10]	Very good 146 (26·6)	Very good 139 (26·6)	1
	Good 232 (42·3)	Good 231 (44·3)	
	Fair 134 (24·4)	Fair 120 (23)	
	Bad 32 (5·8)	Bad 27 (5·2)	
	Very bad 5 (0·9)	Very bad 5 (1)	
Mental age (years) [5, 8]**, mean (SD)**	5.5 (3·0)	4·8 (3)	<·001
Developmental Quotient[5, 8]**, mean (SD)**	0.6 (0·2)	0·5 (0·3)	<·001
ABAS-3[6, 8]**, mean (SD)**	66.4 (13·7)	64·2 (13·1)	·06
SDQ Total score[1]**, mean (SD)**	22.7 (6·5)	18·5 (6·5)	<·001
DAWBA[7, 9]			
Emotional disorders	78 (14·8)	40 (8·1)	<·001
- **Anxiety**	77 (14·7)	40 (8·1%)	<·001
- **Depression**	5 (1)	1 (0·2)	·12
Behavioural disorders	101 (19·1)	49 (10)	<·001
- **Oppositional defiant disorder**	96 (18·1)	48 (9·8)	<·001
- **Conduct disorder**	13 (2·5)	4 (0·8)	·04
Hyperactivity disorder	145 (27·4)	69 (14)	<·001
Autism Spectrum disorder	242 (45·7)	141 (28·7)	<·001

Table 4: Association between SDQ and IMD by variant inheritance in CNV group (n=806)

R^2 = 0·12 for Step 1; R^2 = 0·15 for Step 2; R^2 = 0·16 for Step 3

Model 1 – Associations between SDQ and IMD quintile for individuals by inheritance

Model 2 – Model 1 including confounding variables – child sex, child developmental level as indexed by the developmental quotient (developmental age/chronological age), age of diagnosis and physical health by parent report

Model 3 – Model 2 including interaction factor (inheritance x IMD)

SDQ behaviour difficulties	Model 1			Model 2			Model 3		
	b (SE)	std b	p	b (SE)	std b	p	b (SE)	std b	p
IMD	-0·48 (0·16)	-0·10	·003	-0·34 (0 ·16)	-0·07	·033	-0·47(0 ·23)	-0·10	·036
de novo/familial	4·0 (0·46)	0·31	<·0001	3·7 (0·46)	0·28	<·0001	2·9 (1·06)	0·22	·006
IMD x inheritance	-	-	-	-	-	-	0·26 (0·32)	0·06	·41

References

1.O'Roak BJ, Deriziotis P, Lee C, et al. Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. *Nat Genet* 2011; **43**(6): 585-9.

2.Deciphering Developmental Disorders. Large-scale discovery of novel genetic causes of developmental disorders. *Nature* 2015; **519**(7542): 223-8.

3.Turro E, Astle WJ, Megy K, et al. Whole-genome sequencing of patients with rare diseases in a national health system. *Nature* 2020; **583**(7814): 96-102.

4.Manickam K, McClain MR, Demmer LA, et al. Exome and genome sequencing for pediatric patients with congenital anomalies or intellectual disability: an evidence-based

clinical guideline of the American College of Medical Genetics and Genomics (ACMG).*Genet Med* 2021.

5.Hastings RP, Totsika V, Hayden NK, et al. 1000 Families Study, a UK multiwave cohort investigating the well-being of families of children with intellectual disabilities: cohort profile. *BMJ Open* 2020; **10**(2): e032919.

6.Dunn K, Rydzewska E, MacIntyre C, Rintoul J, Cooper SA. The prevalence and general health status of people with intellectual disabilities and autism co-occurring together:a total population study. *J Intellect Disabil Res* 2019; **63**(4): 277-85.

7.Emerson E, Hatton C. Mental health of children and adolescents with intellectual disabilities in Britain. *Br J Psychiatry* 2007; **191**: 493-9.

8.Ford T, Macdiarmid F, Russell AE, Racey D, Goodman R. The predictors of persistent DSM-IV disorders in 3-year follow-ups of the British Child and Adolescent Mental Health Surveys 1999 and 2004. *Psychol Med* 2017; **47**(6): 1126-37.

9.Altes M, Raymond FL. Avery. 2018. ISBN:9780993146121

10.Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 2015; 559 **17**(5): 405-24.

11.Goodman R, Ford T, Richards H, Gatward R, Meltzer H. The Development and Well-Being Assessment: description and initial validation of an integrated assessment of child and adolescent psychopathology. *J Child Psychol Psychiatry* 2000; **41**(5): 645-55.

12.Goodman A, Heiervang E, Collishaw S, Goodman R. The 'DAWBA bands' as an ordered-categorical measure of child mental health: description and validation in British and Norwegian samples. *Soc Psychiatry Psychiatr Epidemiol* 2011; **46**(6): 521-32.

13.Murray CA, Hastings RP, Totsika V. Clinical utility of the parent-reported Strengths and Difficulties Questionnaire as a screen for emotional and behavioural difficulties in children and adolescents with intellectual disability. *Br J Psychiatry* 2020: 1-3.

14.Allen D, Davies D. Challenging behaviour and psychiatric disorder in intellectual disability. *Curr Opin Psychiatry* 2007; **20**(5): 450-5.

15.DSM-5. *American Psychiatric Association Diagnostic and statistical manual of mental disorders, 5th ed Arlington: American Psychiatric Association, 2013* 2013; **5th**.

16.Goodman R. The Strengths and Difficulties Questionnaire: a research note. *J Child Psychol Psychiatry* 1997; **38**(5): 581-6.

17.Goodman R. Psychometric properties of the strengths and difficulties questionnaire. *J Am Acad Child Adolesc Psychiatry* 2001; **40**(11): 1337-45.

18.Harrison PL, & Oakland, T. . Adaptive behavior assessment system. 2015; **Third** edition. San Antonio, TX: Psychological Corporation.

19.Baker K, Devine RT, Ng-Cordell E, Raymond FL, Consortium IMAGINE-I, Hughes C. Childhood intellectual disability and parents' mental health: integrating social, psychological and genetic influences. *Br J Psychiatry* 2021; **218**(6): 315-322.

20.IBM SPSS Statistics for Windows, Version 24.0. Armonk, NY: IBM Corp.). *IBM Corp Released 2016* 2016; (Version 24.0. Armonk, NY: IBM Corp.).

21.Chawner S, Owen MJ, Holmans P, et al. Genotype-phenotype associations in children with copy number variants associated with high neuropsychiatric risk in the UK (IMAGINE-ID): a case-control cohort study. *Lancet Psychiatry* 2019; **6**(6): 493-505.

22.Harrison PL, and Thomas Oakland. Adaptive behavior assessment system (ABAS-3):Manual. *Pearson Clinical Assessment*, 2015.

23.Satterstrom FK, Kosmicki JA, Wang J, et al. Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. *Cell* 2020; **180**(3): 568-84.

24.Kendall KM, Bracher-Smith M, Fitzpatrick H, et al. Cognitive performance and functional outcomes of carriers of pathogenic copy number variants: analysis of the UK Biobank. *Br J Psychiatry* 2019; **214**(5): 297-304.

25.Reiss F. Socioeconomic inequalities and mental health problems in children and adolescents: a systematic review. *Soc Sci Med* 2013; **90**: 24-31.

26.Robinson EB, St Pourcain B, Anttila V, et al. Genetic risk for autism spectrum disorders and neuropsychiatric variation in the general population. *Nat Genet* 2016; **48**(5): 552-5.

27.Jacquemont S, Coe BP, Hersch M, et al. A higher mutational burden in females supports a "female protective model" in neurodevelopmental disorders. *Am J Hum Genet* 2014; **94**(3): 415-25.

28.Niestroj LM, Perez-Palma E, Howrigan DP, et al. Epilepsy subtype specific copy number burden observed in a genome-wide study of 17,458 subjects. Brain 2020; 143(7):2106-18.