COMMUNICATIONS

Presence of pathogenic copy number variants (CNVs) is correlated with socioeconomic status

ABSTRACT

Socioeconomic status (SES) is a major determinant of health. We studied the Index of Multiple Deprivation Rank of 473 families with individuals with pathogenic autosomal copy number variants (CNVs) and known inheritance status. The IMDR distribution of families with pathogenic CNVs was significantly different from the general population. Families with inherited CNVs were significantly more likely to be living in areas of higher deprivation when compared with families that had individuals with de novo CNVs. These results provide unique insights into biological determinants of SES. As CNVs are relatively frequent in the general population, these results have important medical and policy consequences.

Socioeconomic status (SES) is a measure of an individual's or family's economic and social status based on factors, such as income, education and occupation. SES is a major determinant of health and related outcomes. Lower SES confers increased risk for multifactorial disorders like stroke and cardiovascular disease and plays a key role in child health and development.[1–4] Early life adversity negatively impacts child health and produces lasting and deleterious effects on developmental outcomes.[5] Biological factors including genetic variants that may influence SES are only beginning to be understood. Recently, an association between pathogenic CNVs and lower SES in clinically unaffected adults from the UK Biobank was described.[6] However, this study examined a limited spectrum of common CNVs and there was no information available on the inheritance of these CNVs.[6] The aim of our study was to investigate the correlation of a wide range of unselected pathogenic and likely pathogenic CNVs, and their inheritance pattern, with SES.

We interrogated an anonymised departmental database of results from >17 000 postnatal, mostly paediatric (98%), clinical array comparative genomic hybridisations performed at the Manchester Centre for Genomic Medicine between 2010 and 2017. This database included information on each identified CNV, its clinical pathogenic classification, size, type (loss or gain), where available inheritance status (de novo or inherited from a parent), and postcode of the patient (online supplementary methods).

We identified 473 unique unrelated cases of pathogenic or likely pathogenic autosomal CNVs (table 1 and online supplementary table S1) with known inheritance status ($n_{inherited}$=218 ; $n_{de\ novo}$=255) and available postcode information (online supplementary methods). We then obtained Index of Multiple Deprivation Rank (IMDR) (and the seven constituent domains) associated with each of these postcodes (https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015). English Indices of

Table 1 Frequency of pathogenic CNVs identified in this study

Cytogenetic location of CNVs	De novo loss	De novo gain	Inherited loss	Inherited gain	Total
16p11.2	12	2	18	17	49
1q21.1q21.2	5	1	13	21	40
22q11.21	30	2	3	0	35
15q13.2q13.3	0	0	27	1	28
15q11.2q13.1	11	2	1	6	20
16p13.11	3	0	8	7	18
17q12	4	2	1	8	15
7q11.23	7	4	1	1	13
16p12.2	1	0	12	0	13
17p12	1	0	6	4	11
15q13.3	0	0	8	1	9
17p11.2	5	3	0	1	9
3q29	3	0	0	5	8
2p16.3	3	0	4	0	7
17p13.3	2	3	1	0	6
22q13.33	5	0	1	0	6
21q11.2q22.3	0	4	0	0	4
4p16.3	1	0	2	0	3
6q27	2	0	1	0	3
9p24.3p23	3	0	0	0	3
16p13.11p12.3	1	0	2	0	3
20p13	1	0	1	1	3
22q11.21q11.22	3	0	0	0	3
22q11.23q13.33	0	3	0	0	3
2q37.1q37.3	1	1	0	0	2
5q14.3	2	0	0	0	2
6p25.3p25.2	1	0	1	0	2
7p22.1	2	0	0	0	2
8p23.1	1	0	0	1	2
9p24.2p23	1	0	1	0	2
9p24.3p13.1	0	2	0	0	2
10q26.2q26.3	2	0	0	0	2
10q26.3	1	0	1	0	2
11q24.2q25	2	0	0	0	2
13q12.3q13.1	2	0	0	0	2
15q11.1q13.1	1	1	0	0	2
16p13.3	1	0	1	0	2
17q21.31	2	0	0	0	2
18p11.32p11.21	0	2	0	0	2
18p11.32p11.31	0	0	1	1	2
22q11.1q11.21	0	2	0	0	2
Other CNVs with single instance only	71	28	23	5	127

Pathogenic CNVs identified in this study are listed in the descending order of their total frequencies in our cohort. Note that the individual CNVs grouped to generate frequencies have been grouped according to their chromosomal location and may include overlapping CNVs with different breakpoints. Only chromosomal locations with at least two CNVs in the study have been individually listed here. Full details of all the CNVs and their specific locations are provided in online supplementary table S1.

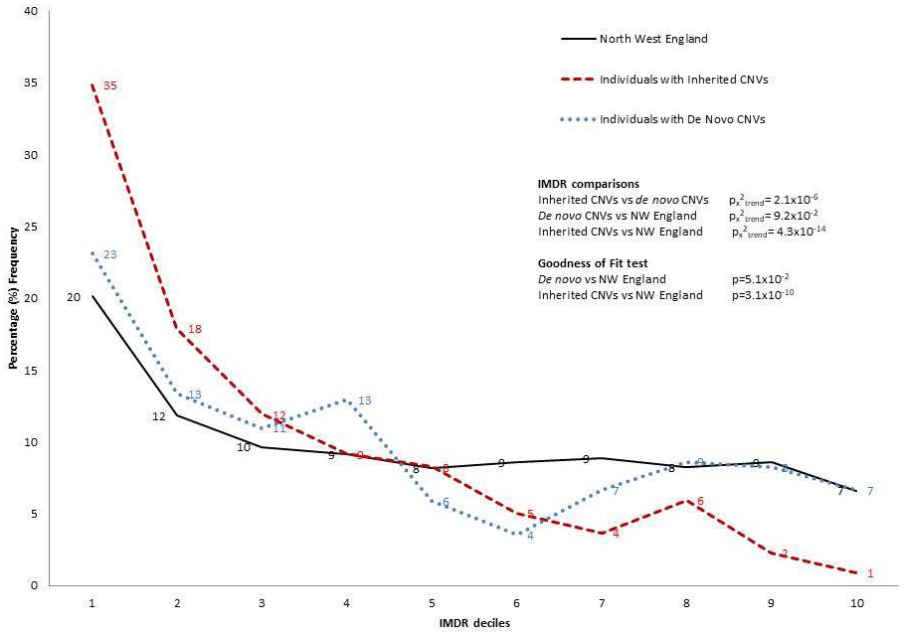

Figure 1 Presence of pathogenic or likely pathogenic CNVs is correlated with socioeconomic status and is primarily driven by inherited CNVs. An anonymised Departmental database of results from over 17 000 postnatal clinical array comparative genomic hybridisation testing performed at Manchester Centre for Genomic Medicine between 2010 and 2017 was curated for CNVs, clinical classification, size, loss or gain status, the inheritance status (de novo or inherited from a parent) where available, and postcode. From the postcodes associated with the CNVs, we retrieved the Index of Multiple Deprivation Rank (IMDR) and its seven constituent domains, using the English indices of deprivation 2015. The distribution of IMDR deciles was compared for inherited (dashed line) versus de novo (dotted line) CNVs against the general population of the North West of England. χ^2 tests of independence and trend were performed on the IMDR deciles. The significance value for the χ^2 test was set at p<0.05.

Deprivation is a widely used measure of SES in health research[7 8] and reflects the SES of the households of the affected individuals in this study (online supplementary table S1).

We found that IMDR composition for families with pathogenic or likely pathogenic CNVs was significantly different when compared with the IMDR of the general population of the North West England ($p_{x\ goodness\ of\ fit}^2 = 1.8\times10^{-8}$). Furthermore, families with inherited pathogenic and likely pathogenic CNVs were significantly more likely to be living in areas of higher deprivation when compared with families that have individuals with de novo pathogenic and likely pathogenic CNVs ($p_{x\ trend}^2 = 2.1\times10^{-6}$) or with the general population of North West England ($p_{x\ trend}^2 = 4.3\times10^{-14}$) (figure 1 and online supplementary figure S1) (online supplementary results). This difference was significant across the following deprivation domains of IMDR—income; employment; health; education, skills and training (online supplementary figures S2–S6). There were no significant differences in the following domains: barriers to housing and services, and living environment (online supplementary figures S7 and S8). We also performed the Jonckheere-Terpstra test in both de novo and inherited CNVs to see if there was any correlation between SES and age of diagnosis. No significant trends were identified (p-value=0.0615 for de novo CNVs and p-value=0.1615 for inherited CNVs) (online supplementary figure S9). IMDR comparisons according to CNV type did not reveal any significant difference between CNV losses and gains (p=0.52) (online supplementary results) (online supplementary figure S10). There was no evidence of an effect of CNV size on IMDR (online supplementary results) (p-value Kruskal-Wallis rank-sum test=0.48) (online supplementary figure S11). Our data show that the presence of pathogenic and likely pathogenic CNVs is correlated with SES. Notably in our cohort, this correlation seems to be driven by partially penetrant inherited CNVs of low effect size rather than de novo CNVs that are usually associated with more severe and more penetrant phenotypes.[9]

The vast majority of the probands in our cohort are children (98%), and therefore, the IMDR dataset reflects the SES of the household in which they are growing up. The parents of children with de novo CNV do not carry the CNV and are not affected by the condition of their children. On the other hand, at least one of the parents of individuals with inherited CNVs will be carrying the same CNV. Almost 51% (n=112) of inherited CNVs in our cohort can be classed as recurrent (table 1 and online supplementary table S1).[9] The penetrance of these recurrent CNVs has been estimated to range between 10% and 62%.[9] Based on reported estimated penetrance of these recurrent CNVs, majority of the carrier parents of our index cases are likely to be classed as medically unaffected. The high level of deprivation observed in our cohort suggests that being a carrier of a low penetrant negatively impacts SES even in the absence of a medical phenotype. Our observations indicate that there are likely subclinical effects in individuals who are medically non-penetrant carriers of milder pathogenic and likely pathogenic CNVs. This agrees with the recent findings from the UK Biobank study whereby carriers of pathogenic CNVs had lower levels of household income and higher deprivation in the absence of neurodevelopmental disorders.[6] Lower SES in families with medically relevant inherited pathogenic and likely pathogenic CNVs with milder phenotype could therefore be due to cumulative multigenerational consequences of these subclinical effects.

These results demonstrate that in addition to the primary phenotypes of the pathogenic and likely pathogenic CNVs, their secondary socioeconomic and resultant medical consequences need to be studied, especially in families with inherited CNVs. The combined frequency for a subset of these pathogenic CNVs in the general population is estimated to be at least 3.8%[10] and therefore are significant in the context of public health. The correlation of SES with inheritance patterns of pathogenic and likely pathogenic CNVs, therefore, provides further unique insights into biological determinants of SES and has important implications for the planning of medical and social services.

George J Burghel,[1,2,1,3,4,1,5] **Unzela Khan**,[1,2,1,3,4,1,5] **Wei-Yu Lin**,[1,2,1,3,4,1,5] **William Whittaker**,[1,2,1,3,4,1,5] **Siddharth Banka**[1,2,1,3,4,1,5]

[1]Manchester Centre for Genomic Medicine, St Mary's Hospital, Manchester University NHS Foundation Trust, Manchester, UK
[2]School of Health Science, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK
[3]Northern Institute for Cancer Research, Medical School, Newcastle University, Newcastle upon Tyne, UK

[4]Division of Population Health, Health Services Research & Primary Care, University of Manchester, Manchester, UK
[5]Division of Evolution and Genomics, School of Biological Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK

Correspondence to Dr Siddharth Banka, Department of Genetic Medicine, University of Manchester, Manchester M13 9WL, UK; Siddharth.Banka@manchester.ac.uk

Contributors SB conceived the project. SB and GJB designed the study and wrote the manuscript. GJB and UK collected and analysed the CNV data. W-YL and WW did the statistical analysis.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/jmedgenet-2019-106292).

To cite Burghel GJ, Khan U, Lin W-Y, *et al*. *J Med Genet* 2020;**57**:70–72.

Received 21 May 2019
Revised 18 July 2019
Accepted 24 July 2019
Published Online First 23 August 2019

J Med Genet 2020;**57**:70–72.
doi:10.1136/jmedgenet-2019-106292

ORCID iD
George J Burghel http://orcid.org/0000-0001-9360-8194

REFERENCES

1. Clark AM, DesMeules M, Luo W, Duncan AS, Wielgosz A. Socioeconomic status and cardiovascular disease: risks and implications for care. *Nat Rev Cardiol* 2009;6:712–22.
2. Cox AM, McKevitt C, Rudd AG, Wolfe CDA. Socioeconomic status and stroke. *Lancet Neurol* 2006;5:181–8.
3. Bradley RH, Corwyn RF. Socioeconomic status and child development. *Annu Rev Psychol* 2002;53:371–99.
4. Hanscombe KB, Trzaskowski M, Haworth CMA, Davis OSP, Dale PS, Plomin R. Socioeconomic status (Ses) and children's intelligence (IQ): in a UK-representative sample Ses moderates the environmental, not genetic, effect on IQ. *PLoS One* 2012;7:e30320.
5. Hackman DA, Farah MJ. Socioeconomic status and the developing brain. *Trends Cogn Sci* 2009;13:65–73.
6. Kendall KM, Bracher-Smith M, Fitzpatrick H, Lynham A, Rees E, Escott-Price V, Owen MJ, O'Donovan MC, Walters JT, Kirov G. Cognitive performance and functional outcomes of carriers of pathogenic copy number variants: analysis of the UK Biobank. *Br J Psychiatry* 2019:1–8.
7. Lang S-J, Abel GA, Mant J, Mullis R. Impact of socioeconomic deprivation on screening for cardiovascular disease risk in a primary prevention population: a cross-sectional study. *BMJ Open* 2016;6:e009984.
8. White J, Rehkopf D, Mortensen LH. Trends in socioeconomic inequalities in body mass index, underweight and obesity among English children, 2007-2008 to 2011-2012. *PLoS One* 2016;11:e0147614.
9. Rosenfeld JA, Coe BP, Eichler EE, Cuckle H, Shaffer LG. Estimates of penetrance for recurrent pathogenic copy-number variations. *Genet Med* 2013;15:478–81.
10. Crawford K, Bracher-Smith M, Owen D, Kendall KM, Rees E, Pardiñas AF, Einon M, Escott-Price V, Walters JTR, O'Donovan MC, Owen MJ, Kirov G. Medical consequences of pathogenic CNVs in adults: analysis of the UK Biobank. *J Med Genet* 2019;56:131–8.