Received: 25 January 2023 | Accepted: 15 July 2023

DOI: 10.1002/aur.2999

SHORT REPORT

Linkage of whole genome sequencing and administrative health data in autism: A proof of concept study

Danielle A. Baribeau[1,2,3] | **Jasleen Arneja**[4] | **Xuesong Wang**[5] | **Jennifer Howe**[6] | **John R. McLaughlin**[7] | **Karen Tu**[8,9] | **Jun Guan**[5] | **Alana Iaboni**[1] | **Elizabeth Kelley**[10] | **Muhammad Ayub**[9,11] | **Robert Nicolson**[12] | **Stelios Georgiades**[13] | **Stephen W. Scherer**[6,14] | **Susan E. Bronskill**[5,7,15,16] | **Evdokia Anagnostou**[1] | **Jennifer D. Brooks**[7]

[1]Bloorview Research Institute, Holland Bloorview Kids Rehabilitation Hospital, Toronto, Ontario, Canada

[2]Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada

[3]SickKids Research Institute, The Hospital for Sick Children, Toronto, Ontario, Canada

[4]Department of Epidemiology Biostatistics and Occupational Health, McGill University, Montreal, Quebec, Canada

[5]ICES, Toronto, Ontario, Canada

[6]The Centre for Applied Genomics, Department of Genetics and Genome Biology, The Hospital for Sick Children, Toronto, Ontario, Canada

[7]Dalla Lana School of Public Health, University of Toronto, Ontario, Canada

[8]North York General Hospital and Toronto Western Family Health Team, University Health Network, Toronto, Ontario, Canada

[9]Department of Family and Community Medicine, University of Toronto, Toronto, Ontario, Canada

[10]Department of Psychiatry, Queens University, Kingston, Ontario, Canada

[11]University College London, London, UK

[12]Department of Psychiatry, Western University, London, Ontario, Canada

[13]Department of Psychiatry and Behavioural Neurosciences, McMaster University, Hamilton, Ontario, Canada

[14]Department of Molecular Genetics and McLaughlin Centre, University of Toronto, Toronto, Ontario, Canada

[15]Women's College Research Institute, Toronto, Ontario, Canada

[16]Hurvitz Brain Sciences Program & Evaluative Clinical Sciences, Sunnybrook Research Institute, Toronto, Ontario, Canada

Correspondence
Danielle A. Baribeau, Bloorview Research Institute, Holland Bloorview Kids Rehabilitation Hospital, Toronto, ON, Canada.
Email: dbaribeau@hollandbloorview.ca

Funding information
Autism Speaks; McLaughlin Centre, Grant/Award Number: MC-2017-04; Ontario Brain Institute

Abstract

Whether genetic testing in autism can help understand longitudinal health outcomes and health service needs is unclear. The objective of this study was to determine whether carrying an autism-associated rare genetic variant is associated with differences in health system utilization by autistic children and youth. This retrospective cohort study examined 415 autistic children/youth who underwent genome sequencing and data collection through a translational neuroscience program (Province of Ontario Neurodevelopmental Disorders Network). Participant data were linked to provincial health administrative databases to identify historical health service utilization, health care costs, and complex chronic medical conditions during a 3-year period. Health administrative data were compared between participants with and without a rare genetic variant in at least 1 of 74 genes associated with autism. Participants with a rare variant impacting an autism-associated gene ($n = 83$, 20%) were less likely to have received psychiatric care (at least one psychiatrist visit: 19.3% vs. 34.3%, $p = 0.01$; outpatient mental health visit: 66% vs. 77%, $p = 0.04$). Health care costs were similar between

groups (median: $5589 vs. $4938, $p = 0.4$) and genetic status was not associated with odds of being a high-cost participant (top 20%) in this cohort. There were no differences in the proportion with complex chronic medical conditions between those with and without an autism-associated genetic variant. Our study highlights the feasibility and potential value of genomic and health system data linkage to understand health service needs, disparities, and health trajectories in individuals with neurodevelopmental conditions.

Lay Summary

In this study, we linked genome-sequencing data from 415 autistic children and youth to their real-world health data. We examined their types of medical visits, health care costs, and number of complex medical problems. We found that autistic children with an underlying genetic diagnosis explaining their autism were less likely to have received mental health care than children without a genetic diagnosis.

KEYWORDS

administrative data, autism, epidemiology, genome sequencing

INTRODUCTION

Autism Spectrum Disorder (herein "autism") is a neurodevelopmental condition that affects 1%–3% of children (Zablotsky et al., 2019). In addition to differences in social development and restricted/repetitive patterns of behavior, autism is associated with higher rates of co-occurring physical health conditions (e.g., epilepsy, gastrointestinal disturbances; Rydzewska et al., 2021), as well as mental health concerns (e.g., anxiety and depression; Lai et al., 2019).

Clinical practice guidelines have recommended genome-wide chromosomal microarray analysis to investigate developmental delays, including autism, for over a decade (Anagnostou et al., 2014). Higher-resolution genetic technologies (including exome and genome sequencing) have recently been recommended for first-line clinical use (Manickam et al., 2021; Srivastava et al., 2019). An autism-associated rare genetic variant (i.e., a genetic diagnosis) can now be identified for approximately 15%–20% of autistic individuals (Srivastava et al., 2019; Stefanski et al., 2021a).

Knowledge of a genetic diagnosis can direct screening for associated health conditions, while contributing to an understanding of underlying mechanisms. Some data suggest that among autistic children, those with an autism-associated genetic diagnosis may have increased rates of developmental delays, intellectual disability, epilepsy, and congenital anomalies compared with those without (Mahjani et al., 2021; Satterstrom et al., 2020).

Despite the above, genetic testing uptake in autism has been low, especially for autistic adolescents and adults (Moreno-De-Luca et al., 2020). With respect to anticipating future physical as well as mental health concerns, the clinical utility of genetic testing in autism remains unclear. One challenge is that most genetic studies in autism are cross-sectional, and the lifelong impact of these genetic changes is unknown for most variants. Linkage of genomic data to health administrative data (including surgeries, emergency department visits, hospital visits, and comorbidities documented in health encounters) could be useful in informing our understanding of genetic differences in health service resource needs/access, and health trajectories in autism over time (Croen et al., 2006).

The broad objective of this study was to demonstrate the feasibility and potential of linking research and health administrative data in autism, to characterize the impact of a genetic diagnosis for autistic children and youth. Specifically, among a research cohort of children with a confirmed diagnosis of autism and who underwent whole genome sequencing, our first aim was to establish secure data linkage to population-based health administrative data. Then we examined whether carrying an autism-associated rare genetic variant identified on genome sequencing was associated with differences in (1) historical health system utilization, (2) historical health system costs, and (3) number of complex chronic medical conditions. We hypothesized that those with an autism-associated genetic diagnosis would have higher health system utilization, including higher overall health care costs, and a greater number of complex chronic conditions.

METHODS

This retrospective cohort study linked clinical and genetic data from the Province of Ontario Neurodevelopmental Disorders (POND) Network, (a large translational neuroscience research program), with population-based health administrative data to study historical health service utilization in a cohort of autistic children.

Data sources

POND

The POND Network (https://pond-network.ca/) is a multisite integrated discovery program of children and youth which aims to understand the neurobiology of neurodevelopmental disorders, including autism ($n = \sim$3000 in 2023). POND participants were recruited from one of five clinical-research sites. Inclusion criteria were broad: a neurodevelopmental disorder diagnosis (i.e., autism, attention-deficit/hyperactivity disorder [ADHD], or other), and birth after 35 weeks gestation. At intake, POND participants underwent comprehensive clinical phenotyping on numerous validated measures (described by Baribeau et al., 2015). Autism diagnoses were supported using the Autism Diagnostic Observation Schedule (Lord et al., 2000) and the Autism Diagnostic Interview—Revised (Lord et al., 1994). Participants also provided a genetic sample. Recruitment began in 2011 and will continue through 2026.

Genome sequencing was carried out at The Centre for Applied Genomics (TCAG) at SickKids, using Illumina platforms (HiSeq 2000, HiSeq 2500, or HiSeq X) and variant calls were made using TCAG established pipelines (Trost et al., 2022; Yuen RK et al., 2017).

Provincial administrative health data

Children and youth living in Ontario, Canada, are insured under a single-payer health system that covers physician and hospital services and procedures. Administrative health data are generated for each encounter. We linked health administrative data to POND participant data using unique encoded identifiers and analyzed this cohort at ICES (formerly known as the Institute for Clinical Evaluative Sciences). ICES is an independent, non-profit research institute whose legal status under Ontario's health information privacy law allows it to collect and analyze health care and demographic data for health system evaluation and improvement. Databases queried through ICES (see Table S1) provided information on hospital visits, outpatient clinics, day surgeries, emergency department visits, as well as diagnoses and medical service billings by Ontario physicians. The participants' residential postal codes were used to determine neighborhood-level income quintile (range 1 [lowest] to 5 [highest]) and rurality (community with <10,000 residents = rural; 10,000–100,000 = suburban; > 100,000 = urban) using Statistics Canada postal code conversion file plus software (PCCF+) (Kapral et al., 2012).

Study cohort

We selected all autistic POND participants recruited prior to 2018, who had consented for their data to be linked to province-wide health administrative databases ($n = 419$). The POND recruitment date (March 2012 to September 2017) was the participant index date. All data reported in the tables were collapsed to cell sizes greater than five to minimize the risk of participant re-identification. Participant clinical and demographic characteristics included age, sex, ADHD symptoms, autistic traits, adaptive skills, and intelligence quotient (IQ), assessed on standard measures, including the Strengths and Weaknesses of ADHD-symptoms and Normal-behavior (SWAN) rating scale, the Social Communication Questionnaire (SCQ), the Adaptive Behavior Assessment System II, and IQ measures (Baribeau et al., 2015).

For this feasibility study, 74 genes (Table S2) with evidence in more than one academic publication reporting rare high-impact variants believed to be contributing to autism were selected to identify individuals with autism-associated variants on the index date. Individuals with loss-of-function variants, copy number variants (all types impacting coding exons), de novo damaging missense variants, and missense variants predicted to be highly damaging that were inherited or of unknown inheritance *in these genes* were identified and categorized as having an autism-associated variant (yes/no). While this does not exclude the presence of impactful developmental variants (intronic, non-coding, etc.,) in these genes, or other variants in other relevant genes among individuals in the "no variant" group, the number of individuals with identified variants ($n = 83/415$, 20%) was comparable to contemporary estimates using genome-wide variant calls (e.g., 17%; Stefanski et al., 2021b). This gene list was therefore considered a fair proxy for dividing the sample by variant status.

Statistical analyses

A look-back window of 3 years from the index date was used to identify historical health system utilization, including the number and types of specialist visits, visit setting (outpatient, hospital, emergency), as well as visits for a mental health indication (Mental Health and Addictions Program Framework Research Team, 2021). We also calculated the total healthcare costs per POND participant over the 3-year period across all healthcare settings paid for by the province in 2020 equivalent Canadian dollars based on a previously published costing algorithm (Wodchis et al., 2013). High-cost POND participants were defined as individuals within the top 20% of the study cohort.

We identified lifetime history of one or more complex chronic medical conditions on the index date. These were defined as any medical condition that can be reasonably expected to last at least 12 months and to involve either several different organ systems or one organ system severely enough to require specialty pediatric care.

TABLE 1 Demographic and clinical characteristics of $N = 415$ autistic POND participants with and without an autism-associated rare genetic variant identified by genome sequencing, in Ontario, Canada.

Variable	Variant ($n = 83$)	No variant[a] ($n = 332$)	*p*-value	Standardized difference
Age				
Mean ± SD	9.17 ± 4.50	9.52 ± 4.81	0.5	0.08
Median and IQR	9 (5–12)	10 (5–13)	0.6	0.06
Female sex	22 (26.5%)	75 (22.6%)	0.5	0.09
Neighborhood income quintile				
1st (lowest)	1–5 (1.2–6.0%)	52 (15.7%)	0.1	**0.31**
2nd	15–19 (18.1–22.9%)	63 (19.0%)		0.02
3rd	17 (20.5%)	47 (14.2%)		**0.17**
4th	19 (22.9%)	81 (24.4%)		0.04
5th (highest)	27 (32.5%)	89 (26.8%)		**0.13**
Rurality				
Urban/suburban	78–82 (94–99%)	316 (95.2%)	0.3	**0.13**
Rural	1–5 (1.2–6.0%)	16 (4.8%)		**0.13**
Full-scale IQ				
Mean ± SD	80.62 ± 27.49	87.50 ± 24.45	0.06	**0.16**
Median (IQR)	85.0 (52.0–104.0)	92.0 (70.0–106.0)	0.1	**0.10**
Adaptive skills on ABAS II, Mean ± SD				
General adaptive composite score	64.71 ± 18.08	63.97 ± 17.01	0.7	0.04
Conceptual composite score	70.87 ± 17.83	70.76 ± 17.46	0.9	0.01
Social composite score	70.60 ± 14.86	69.65 ± 13.82	0.6	0.07
Practical composite score	64.17 ± 19.94	63.13 ± 18.37	0.7	0.06
Elevated autistic traits: (SCQ above clinical cut-off of 15)	52 (62.7%)	191 (57.5%)	0.7	**0.10**
Elevated ADHD symptoms: (SWAN score above clinical cut-off of 6)	28 (33.7%)	111 (33.4%)	0.6	0.01
Elevated Oppositional traits: (Oppositional items on SWAN above 4)	23 (27.7%)	88 (26.5%)	0.9	0.03

Note: Cell sizes less than 6 have been collapsed or are omitted to de-identify patient data. Standardized differences are interpreted like effect sizes (i.e., 0.2, 0.5, and 0.8 can be used to represent small, medium, and large effect sizes, respectively). Those greater than 0.1 are bolded to show small/ potential differences.
Abbreviations: ABAS, adaptive behavior assessment system; ADHD, attention-deficit/hyperactivity disorder; IQ, intelligence quotient on one of several measures (see Baribeau et al., 2015); SCQ, social communication questionnaire, SWAN, strengths and weaknesses of ADHD-symptoms and normal-behavior (SWAN) rating scale.
[a]Variant status: Presence or absence of an underlying rare genetic variant associated with autism.

Complex chronic conditions were identified using a previously published algorithm by Feudtner et al. (2014), based on International Classification of Disease and Related Health Problems, Tenth Revision, Canada (ICD-10-CA) and the Canadian Classification of Health Interventions (CCI) codes, which was modified slightly (Table S3).

To compare demographic variables, clinical characteristics, and health system utilization/costs between individuals with and without an autism-associated genetic variant, the chi-squared test was used for categorical variables, the one-way ANOVA was used for comparing means, and the Kruskal-Wallis test was used for comparing medians (IQR). Uncorrected p-values are reported to convey trends. Multivariable logistic regression was used to estimate odds ratios (OR) and 95% confidence intervals (CI) for the relationship between carrying a genetic variant and being a high-cost POND participant. This model was adjusted for age, sex, neighborhood-level income quintile, and geographic location (rural/suburban/urban).

RESULTS

POND study participants were ethnically diverse, and comparable to Ontario population level census data (Aboriginal 4%, Arab 2%, Black 6%, Chinese 5%, East Asian 1%, South Asian 5%, Southeast Asian 1%, West Asian 1%, Filipino 2%, Japanese 1%, Jewish 8%, Korean 1%, Latin American/Hispanic 5%, White 62%, Other 3%. Note that participants could select multiple ethnicity categories). Of 851 autistic POND participants recruited prior to 2018, 419 (49%) consented to health administrative data linkage; there were no significant differences in age, sex, IQ, adaptive skills, or developmental and behavioral symptoms/traits between those who did and did not consent to data linkage (data not shown).

TABLE 2 Historical health system utilization, costs, and complex chronic conditions, by autistic POND participants with and without an autism-associated rare genetic variant.

		Variant ($n = 83$)	No variant ($n = 332$)	*p*-value	Standardized difference
Health system utilization					
Age at first specialist visit	Mean ± SD	6.7 ± 4.5	7.1 ± 4.8	0.5	0.08
Number of distinct specialists visited	Median (IQR)	4 (2–5)	3 (2–5)	0.6	0.07
Any specialist	*n* (%)	75–78 (90.4–94.0%)	316 (95.2%)	0.7	0.05
Pediatrics	*n* (%)	73 (88.0%)	282 (84.9%)	0.5	0.09
Internal medicine	*n* (%)	7 (8.4%)	32 (9.6%)	0.7	0.04
ENT	*n* (%)	32 (38.6%)	97 (29.2%)	0.1	**0.20**
Dermatology	*n* (%)	9 (10.8%)	27 (8.1%)	0.4	0.09
Emergency medicine	*n* (%)	8 (9.6%)	26 (7.8%)	0.6	0.06
Psychiatry	*n* (%)	**16 (19.3%)**	**114 (34.3%)**	**0.01**	**0.35**
Cardiology	*n* (%)	6 (7.2%)	29 (8.7%)	0.7	0.06
Ophthalmology	*n* (%)	13 (15.7%)	44 (13.3%)	0.6	0.07
Neurology	*n* (%)	20 (24.1%)	55 (16.6%)	0.1	**0.19**
Any surgical visit	*n* (%)	26 (31.3%)	94 (28.3%)	0.6	0.07
Orthopedic surgery	*n* (%)	11 (13.3%)	31 (9.3%)	0.3	**0.12**
Same day surgery	*n* (%)	12 (14.5%)	56 (16.9%)	0.6	0.07
Any emergency department visit	*n* (%)	43 (51.8%)	193 (58.1%)	0.3	**0.13**
Emergency department visit (psychiatric)	*n* (%)	1–5 (1–6%)	19 (5.7%)	0.2	**0.17**
Emergency department visit (non-psychiatric)	*n* (%)	42 (50.6%)	189 (56.9%)	0.3	0.13
Mental health—outpatient	***n* (%)**	**55 (66.3%)**	**256 (77.1%)**	**0.04**	**0.24**
Any hospitalization	*n* (%)	14 (16.9%)	41 (12.3%)	0.3	**0.13**
Psychiatric hospitalization	*n* (%)	0 (0.0%)	7 (2.1%)	0.2	**0.21**
Health care costs in 2020 Canadian Dollars					
High-cost participant	*n* (%)	19(22.9%)	64(19.3%)	0.5	0.09
Total health care costs	Mean (± SD)	$8595.39 (± 9093.70)	$9905.02 (± 22,580.14)	0.6	0.07
	Median (IQR)	$5589.00 (2620.00-10,346.00)	$4938.00 (2372.50-8792.00)	0.4	**0.10**
Hospitalization costs	Mean (± SD)	1253.28 ± 3405.32	1822.73 ± 9536.29	0.6	0.08
	Median (IQR)	0.00 (0.00–0.00)	0.00 (0.00–0.00)	0.3	**0.13**
Emergency department costs	Mean ± SD	$289.34 ± 479.96	$309.20 ± 568.29	0.8	0.04
	Median (IQR)	$96.00 (0.00–475.00)	$124.00 (0.00–357.00)	0.6	0.07
Physician billing costs	Mean ± SD	$2937.34 ± 2467.74	$2972.44 ± 3160.45	0.9	0.01
	Median (IQR)	$2314.00 (1142.00-4181.00)	$2069.00 (1053.50-3941.50)	0.6	0.06
Publicly funded medication costs	Mean ± SD	$614.70 ± 3619.99	$461.48 ± 1982.68	0.6	0.05
	Median (IQR)	$0.00 (0.00–61.00)	$0.00 (0.00–14.00)	0.2	**0.14**
Home care service costs	Mean ± SD	$1780.57 ± 3735.44	$2428.45 ± 16,034.16	0.7	0.06
	Median (IQR)	$0.00 (0.00–2262.00)	$0.00 (0.00–1761.50)	0.9	0.01
Number with one or more complex chronic medical conditions	*n* (%)	8 (9.6%)	28 (8.4%)	0.7	0.04

Note: Cell sizes less than six have been collapsed or are omitted to de-identify patient data. Total health care costs include all costs for publicly funded health services. Both health system utilization and costs were calculated from a 3-year window. Complex chronic medical conditions were identified from lifetime health administrative data. Standardized differences are interpreted like effect sizes (i.e., 0.2, 0.5, and 0.8 can be used to represent small, medium, and large effect sizes, respectively). Those greater than 0.1 are bolded to show small/ potential differences.

Research and health administrative data linkage through dual-end encryption of individual health card numbers was successful for 415 of 419 (99%) of participants.

There were no statistically significant differences in age, sex, socioeconomic, or clinical characteristics between autistic participants with and without one or more rare high-impact variants impacting an autism-associated gene (Table 1).

Regarding historical health system utilization, mental health visits were the only service that differed between the genomic groups (Table 2). Autistic children with an autism-associated variant were less likely to have received psychiatric care than those without, including outpatient visits for mental health (66.3% vs. 77.1%, $p = 0.04$) and care from a psychiatrist (19.3% vs. 34.3%, $p = 0.01$).

Costs were not significantly different between groups (Table 2). In the multivariable logistic regression model, the adjusted estimate did not identify variant status as being significantly associated with odds of being a high-cost POND participant (OR 1.36, 95% CI 0.75 to 2.47, $p = 0.3$) (Table S4).

There were no significant differences in the number of complex chronic conditions between participants with and without a variant impacting an autism-associated gene (Table 2).

Results were unchanged after excluding four participants with other recurrent CNVs that did not contain a known autism gene as of 2020.

DISCUSSION

In this feasibility study, within a group of school-aged autistic children and youth, we successfully linked research-based genome sequencing data to longitudinal real-word evidence from health administrative data. In line with our broad objective, this study demonstrated the potential utility of this linkage to examine health trajectories in rare genetic and/or neurodevelopmental disorders, without incurring additional participant burden, and while circumventing bias to due recall or unequal loss-to-follow up.

Overall, we did not detect a significant difference between those with and without a high-impact genetic variant in an autism-associated gene regarding health system costs or number of complex chronic conditions. Unexpectedly, we found that variant carriers may have lower rates of contact with the mental health care system.

The lower rates of psychiatric health care utilization by those with autism-associated variants are notable. This could suggest differences in autism ascertainment pathways (medical vs. psychiatric assessments), diagnostic overshadowing by the genetic diagnosis and/or intellectual disability, or systemic barriers to receiving psychiatric care for those with rare genetic disorders. Data could also support distinct genetic mechanisms towards psychiatric vulnerability in autism. For example, emerging data suggest there is: (1) an inverse correlation between high-impact de novo variants and autism polygenic scores—(i.e., genetic contributors to autism seem to include *either* high polygenic likelihood *OR* high-impact rare variants; Antaki et al., 2022; Warrier et al., 2022), (2) there are general effects of autism polygenic scores on psychopathology broadly (i.e., high polygenic scores may increase likelihood for developmental and mental health differences; Kember et al., 2021; Neumann et al., 2022), and (3) that in neuropsychiatric disorders, individuals with inherited genetic risk factors may be more likely to experience socioeconomic/environmental adversity (Ratanatharathorn et al., 2021; Wolstencroft et al., 2022). Together, these findings set the stage for future studies examining the types of psychiatric conditions, service needs, and barriers faced by those with genetic forms of autism, as well as the relative contributions of environmental factors, comorbid medical conditions, and rare/common and inherited/de novo variants, to help anticipate mental health care needs in autistic youth.

The current cohort is well characterized, with gold-standard autism diagnostic procedures. Our study examined comorbidities with currently validated administrative health data algorithms for case identification in children within Ontario provincial administrative data. While examining the number of comorbidities is a strength of our analysis, we were not able to examine potentially important comorbidities individually (e.g., epilepsy, gastrointestinal issues). We were also not able to examine services administered by other health professionals (e.g., psychologists, social workers, occupational and physical therapists) not captured by physician billing codes in administrative health data, therefore findings likely underestimate total health care as well as mental-health care utilization. We were also underpowered to detect small differences between groups. The gene-based variant classification approach was established at study start and was not as comprehensive as more contemporary pipelines. Misclassification would have biased away from detecting differences between groups. Differences would be non-significant after a correction for multiple comparisons. Given the exploratory nature of this work, findings need to be further examined in larger samples and with longer follow-up. In addition, participants underwent genome sequencing through a research study; we did not examine the cost-effectiveness of genome sequencing in autism specifically. High-cost participant status was estimated relative to other study participants, not the general population.

In summary, this preliminary analysis highlights the feasibility and potential of genomic and health system data linkage. Given (1) the high prevalence of co-occurring psychiatric disorders and mental health service utilization by autistic youth (Martini et al., 2022),

(2) growing access to clinical genome sequencing across medical settings (Costain et al., 2021), and (3) resulting anxiety and uncertainty re: developmental and psychiatric implications of a rare genetic diagnosis (Perlman et al., 2023), studies like ours set the grounds for the exploration of the clinical utility of using genetic testing data to understand and anticipate clinical mental health risk and care needs in neurodevelopmental and rare genetic disorders.

ACKNOWLEDGMENTS

We acknowledge the resources from Autism Speaks-MSSNG (https://research.mss.ng/). This study was supported by ICES, which is funded by an annual grant from the Ontario Ministry of Health (MOH) and the Ministry of Long-Term Care (MLTC). This study also received funding from the University of Toronto McLaughlin Centre and the Ontario Brain Institute (Accelerator Grant: MC-2017-04). Danielle A. Baribeau acknowledges the Arthur Family Foundation for its generous support of her research time at Holland Bloorview. Karen Tu is supported by a Research Scholar Award from the Department of Family and Community Medicine at the University of Toronto. Stephen W. Scherer holds the Northbridge Chair in Pediatric Research at The Hospital for Sick Children and the University of Toronto. This document used data adapted from the Statistics Canada Postal Code Conversion File, which is based on data licensed from Canada Post Corporation, and/or data adapted from the Ontario Ministry of Health Postal Code Conversion File, which contains data copied under license from ©Canada Post Corporation and Statistics Canada. Parts of this material are based on data and/or information compiled and provided by CIHI and the Ontario Ministry of Health. The analyses, conclusions, opinions, and statements expressed herein are solely those of the authors and do not reflect those of the funding or data sources; no endorsement is intended or should be inferred.

CONFLICT OF INTEREST STATEMENT

Evdokia Anagnostou has received consultation fees from Roche, Quadrant, and Oron; grant funding from Roche; in-kind supports from AMO Pharma and CRR; editorial honoraria from Wiley; and book royalties from APPI and Springer. She co-holds a patent for the device Anxiety Meter (patent # US20160000365A1). Stephen W. Scherer is on the Scientific Advisory Committee of Population Bio; serves as a Highly Cited Academic Advisor for King Abdulaziz University; has intellectual property from aspects of his research held at The Hospital for Sick Children licensed to Athena Diagnostics and Population Bio. Danielle A. Baribeau has received research funds from MapLight therapeutics. These relationships did not influence data interpretation/ presentation study but are disclosed for consideration. The remaining authors declare no conflicts.

DATA AVAILABILITY STATEMENT

The Province of Ontario Neurodevelopmental Disorders (POND) Network data are available by request through open data sharing via https://www.braincode.ca/.

ETHICS STATEMENT

This project was approved by Research Ethics Boards at the University of Toronto and at The Hospital for Sick Children (SickKids), Toronto, Canada.

ORCID

Danielle A. Baribeau https://orcid.org/0000-0003-4461-5851
Jasleen Arneja https://orcid.org/0000-0003-1913-6699
Elizabeth Kelley https://orcid.org/0000-0001-7742-6542

REFERENCES

Anagnostou, E., Zwaigenbaum, L., Szatmari, P., Fombonne, E., Fernandez, B. A., Woodbury-Smith, M., Brian, J., Bryson, S., Smith, I. M., Drmic, I., Buchanan, J. A., Roberts, W., & Scherer, S. W. (2014). Autism spectrum disorder: Advances in evidence-based practice. *CMAJ*, *186*(7), 509–519. https://doi.org/10.1503/cmaj.121756

Antaki, D., Guevara, J., Maihofer, A. X., Klein, M., Gujral, M., Grove, J., Carey, C. E., Hong, O., Arranz, M. J., Hervas, A., Corsello, C., Vaux, K. K., Muotri, A. R., Iakoucheva, L. M., Courchesne, E., Pierce, K., Gleeson, J. G., Robinson, E. B., Nievergelt, C. M., & Sebat, J. (2022). A phenotypic spectrum of autism is attributable to the combined effects of rare variants, polygenic risk and sex. *Nature Genetics*, *54*(9), 1284–1292. https://doi.org/10.1038/s41588-022-01064-5

Baribeau, D. A., Doyle-Thomas, K. A., Dupuis, A., Iaboni, A., Crosbie, J., McGinn, H., Arnold, P. D., Brian, J., Kushki, A., Nicolson, R., Schachar, R. J., Soreni, N., Szatmari, P., & Anagnostou, E. (2015). Examining and comparing social perception abilities across childhood-onset neurodevelopmental disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, *54*(6), 479–486 e1. https://doi.org/10.1016/j.jaac.2015.03.016

Costain, G., Cohn, R. D., Scherer, S. W., & Marshall, C. R. (2021). Genome sequencing as a diagnostic test. *Canadian Medical Association Journal*, *193*(42), E1626–E1629. https://doi.org/10.1503/cmaj.210549

Croen, L. A., Najjar, D. V., Ray, G. T., Lotspeich, L., & Bernal, P. (2006). A comparison of health care utilization and costs of children with and without autism spectrum disorders in a large group-model health plan. *Pediatrics*, *118*(4), e1203–e1211. (In English). https://doi.org/10.1542/peds.2006-0127

Feudtner, C., Feinstein, J. A., Zhong, W., Hall, M., & Dai, D. (2014). Pediatric complex chronic conditions classification system version 2: Updated for ICD-10 and complex medical technology dependence and transplantation. *BMC Pediatrics*, *14*, 199. (In English). https://doi.org/10.1186/1471-2431-14-199

Kapral, M. K., Fang, J., Chan, C., Alter, D. A., Bronskill, S. E., Hill, M. D., Manuel, D. G., Tu, J. V., & Anderson, G. M. (2012). Neighborhood income and stroke care and outcomes. *Neurology*, *79*(12), 1200–1207.

Kember, R. L., Merikangas, A. K., Verma, S. S., Verma, A., Judy, R., Damrauer, S. M., Ritchie, M. D., Rader, D. J., Bućan, M., Abecasis, G., Baras, A., Cantor, M., Coppola, G., Economides, A., Lotta, L., Overton, J. D., Reid, J. G., Shuldiner, A., Beechert, C., … Mitnaul, L. J. (2021). Polygenic risk of psychiatric disorders exhibits cross-trait associations in electronic health record data from European ancestry individuals. *Biological Psychiatry*, *89*(3), 236–245. https://doi.org/10.1016/j.biopsych.2020.06.026

Lai, M. C., Kassee, C., Besney, R., Bonato, S., Hull, L., Mandy, W., Szatmari, P., & Ameis, S. H. (2019). Prevalence of co-occurring mental health diagnoses in the autism population: A systematic review and meta-analysis. *Lancet Psychiatry*, *6*, 819–829. https://doi.org/10.1016/S2215-0366(19)30289-5

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Jr., Leventhal, B. L., DiLavore, P. C., Pickles, A., & Rutter, M. (2000). The autism diagnostic observation schedule-generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, *30*(3), 205–223.

Lord, C., Rutter, M., & le Couteur, A. (1994). Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, *24*(5), 659–685.

Mahjani, B., de Rubeis, S., Gustavsson Mahjani, C., Mulhern, M., Xu, X., Klei, L., Satterstrom, F. K., Fu, J., Talkowski, M. E., Reichenberg, A., Sandin, S., Hultman, C. M., Grice, D. E., Roeder, K., Devlin, B., & Buxbaum, J. D. (2021). Prevalence and phenotypic impact of rare potentially damaging variants in autism spectrum disorder. *Molecular Autism*, *12*(1), 65. https://doi.org/10.1186/s13229-021-00465-3

Manickam, K., McClain, M. R., Demmer, L. A., Biswas, S., Kearney, H. M., Malinowski, J., Massingham, L. J., Miller, D., Yu, T. W., Hisama, F. M., & ACMG Board of Directors. (2021). 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). *Genetics in Medicine*, *23*, 2029–2037. https://doi.org/10.1038/s41436-021-01242-6

Martini, M. I., Kuja-Halkola, R., Butwicka, A., du Rietz, E., D'Onofrio, B. M., Happé, F., Kanina, A., Larsson, H., Lundström, S., Martin, J., Rosenqvist, M. A., Lichtenstein, P., & Taylor, M. J. (2022). Sex differences in mental health problems and psychiatric hospitalization in autistic Young adults. *JAMA Psychiatry*, *79*, 1188–1198. https://doi.org/10.1001/jamapsychiatry.2022.3475

Mental Health and Addictions Program Framework Research Team. (2021). *Mental health and addictions system performance in Ontario: 2021 scorecard*. ICES.

Moreno-De-Luca, D., Kavanaugh, B. C., Best, C. R., Sheinkopf, S. J., Phornphutkul, C., & Morrow, E. M. (2020). Clinical genetic testing in autism Spectrum disorder in a large community-based population sample. *JAMA Psychiatry*, *77*(9), 979–981. (In English). https://doi.org/10.1001/jamapsychiatry.2020.0950

Neumann, A., Jolicoeur-Martineau, A., Szekely, E., Sallis, H. M., O'Donnel, K., Greenwood, C. M. T., Levitan, R., Meaney, M. J., Wazana, A., Evans, J., & Tiemeier, H. (2022). Combined polygenic risk scores of different psychiatric traits predict general and specific psychopathology in childhood. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *63*(6), 636–645. https://doi.org/10.1111/jcpp.13501

Perlman, P., Vorstman, J., Hoang, N., Summers, J., Baribeau, D., Cunningham, J., & Mulsant, B. H. (2023). Support to caregivers who have received genetic information about neurodevelopmental and psychiatric vulnerability in their young children: A narrative review. *Clinical Genetics*, *104*, 163–173. https://doi.org/10.1111/cge.14349

Ratanatharathorn, A., Koenen, K. C., Chibnik, L. B., Weisskopf, M. G., Rich-Edwards, J. W., & Roberts, A. L. (2021). Polygenic risk for autism, attention-deficit hyperactivity disorder, schizophrenia, major depressive disorder, and neuroticism is associated with the experience of childhood abuse. *Molecular Psychiatry*, *26*(5), 1696–1705. https://doi.org/10.1038/s41380-020-00996-w

Rydzewska, E., Dunn, K., & Cooper, S.-A. (2021). Umbrella systematic review of systematic reviews and meta-analyses on comorbid physical conditions in people with autism spectrum disorder. *The British Journal of Psychiatry*, *218*(1), 10–19. https://doi.org/10.1192/bjp.2020.167

Satterstrom, F. K., Kosmicki, J. A., Wang, J., Breen, M. S., de Rubeis, S., An, J. Y., Peng, M., Collins, R., Grove, J., Klei, L., Stevens, C., Reichert, J., Mulhern, M. S., Artomov, M., Gerges, S., Sheppard, B., Xu, X., Bhaduri, A., Norman, U., … Walters, R. K. (2020). Large-scale exome sequencing study implicates both developmental and functional changes in the neurobiology of autism. *Cell*, *180*(3), 568–584.e23. https://doi.org/10.1016/j.cell.2019.12.036

Srivastava, S., Love-Nichols, J. A., Dies, K. A., Ledbetter, D. H., Martin, C. L., Chung, W. K., Firth, H. V., Frazier, T., Hansen, R. L., Prock, L., Brunner, H., Hoang, N., Scherer, S. W., Sahin, M., Miller, D. T., & NDD Exome Scoping Review Work Group. (2019). Meta-analysis and multidisciplinary consensus statement: Exome sequencing is a first-tier clinical diagnostic test for individuals with neurodevelopmental disorders. *Genetics in Medicine*, *21*, 2413–2421. https://doi.org/10.1038/s41436-019-0554-6

Stefanski, A., Calle-López, Y., Leu, C., Pérez-Palma, E., Pestana-Knight, E., & Lal, D. (2021a). Clinical sequencing yield in epilepsy, autism spectrum disorder, and intellectual disability: A systematic review and meta-analysis. *Epilepsia*, *62*(1), 143–151.

Trost, B., Thiruvahindrapuram, B., Chan, A. J. S., Engchuan, W., Higginbotham, E. J., Howe, J. L., Loureiro, L. O., Reuter, M. S., Roshandel, D., Whitney, J., Zarrei, M., Bookman, M., Somerville, C., Shaath, R., Abdi, M., Aliyev, E., Patel, R. V., Nalpathamkalam, T., Pellecchia, G., … Scherer, S. W. (2022). Genomic architecture of autism from comprehensive whole-genome sequence annotation. *Cell*, *185*(23), 4409–4427.e18. https://doi.org/10.1016/j.cell.2022.10.009

Warrier, V., Zhang, X., Reed, P., Havdahl, A., Moore, T. M., Cliquet, F., Leblond, C. S., Rolland, T., Rosengren, A., EU-AIMS LEAP, Caceres, A. S. J., Hayward, H., Crawley, D., Faulkner, J., Sabet, J., Ellis, C., Oakley, B., Loth, E., Charman, T., … Baron-Cohen, S. (2022). Genetic correlates of phenotypic heterogeneity in autism. *Nature Genetics*, *54*(9), 1293–1304. https://doi.org/10.1038/s41588-022-01072-5

Wodchis, W., Bushmeneva, K., Nikitovic, M., & McKillop, I. (2013). *Guidelines on person-level costing using administrative databases in Ontario*. Health System Performance Research Network. University of Toronto TSpace.

Wolstencroft, J., Wicks, F., Srinivasan, R., Wynn, S., Ford, T., Baker, K., Chawner, S. J. R. A., Hall, J., van den Bree, M. B. M., Owen, M. J., Skuse, D., Raymond, F. L., Wolstencroft, J., Wicks, F., Srinivasan, R., Erwood, M., Lafont, A., Timur, H., Ye, Z., … Raymond, F. L. (2022). Neuropsychiatric risk in children with intellectual disability of genetic origin: IMAGINE, a UK national cohort study. *Lancet Psychiatry*, *9*(9), 715–724. https://doi.org/10.1016/s2215-0366(22)00207-3

Yuen RK, C., Merico, D., Bookman, M., Howe, J. L., Thiruvahindrapuram, B., Patel, R. V., Whitney, J., Deflaux, N., Bingham, J., Wang, Z., Pellecchia, G., Buchanan, J. A., Walker, S., Marshall, C. R., Uddin, M., Zarrei, M., Deneault, E., D'Abate, L., Chan, A. J. S., … Scherer, S. W. (2017). Whole genome sequencing resource identifies 18 new candidate genes for autism spectrum disorder. *Nature Neuroscience*, *20*(4), 602–611. https://doi.org/10.1038/nn.4524

Zablotsky, B., Black, L. I., Maenner, M. J., Schieve, L. A., Danielson, M. L., Bitsko, R. H., Blumberg, S. J., Kogan, M. D., & Boyle, C. A. (2019). Prevalence and trends of developmental disabilities among children in the United States: 2009–2017. *Pediatrics*, *144*(4), e.20190811.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Baribeau, D. A., Arneja, J., Wang, X., Howe, J., McLaughlin, J. R., Tu, K., Guan, J., Iaboni, A., Kelley, E., Ayub, M., Nicolson, R., Georgiades, S., Scherer, S. W., Bronskill, S. E., Anagnostou, E., & Brooks, J. D. (2023). Linkage of whole genome sequencing and administrative health data in autism: A proof of concept study. *Autism Research*, *16*(8), 1600–1608. https://doi.org/10.1002/aur.2999