The Journal of Molecular Diagnostics, Vol. 24, No. 3, March 2022

ELSEVIER

the Journal of Molecular Diagnostics
jmdjournal.org

Molecular Diagnostic Outcomes from 700 Cases

What Can We Learn from a Retrospective Analysis of Clinical Exome Sequencing?

Jill R. Murrell,[*†] Addie May I. Nesbitt,[*] Samuel W. Baker,[*] Kieran B. Pechter,[*] Jorune Balciuniene,[*] Xiaonan Zhao,[*] Elizabeth H. Denenberg,[*] Elizabeth T. DeChene,[*] Chao Wu,[*] Pushkala Jayaraman,[*] Kajia Cao,[*] Michael Gonzalez,[*] Marcella Devoto,[‡§¶] Alessandro Testori,[‡‖] John D. Monos,[*] Matthew C. Dulik,[*†] Laura K. Conlin,[*†] Minjie Luo,[*†] Kristin McDonald Gibson,[*] Qiaoning Guan,[*] Mahdi Sarmady,[*†] Elizabeth Bhoj,[†‡] Ingo Helbig,[**††] Elaine H. Zackai,[‡§] Emma C. Bedoukian,[‡‡] Alisha Wilkens,[*‡‡] Jennifer Tarpinian,[‡‡] Kosuke Izumi,[‡§‡‡] Cara M. Skraban,[‡§‡‡] Matthew A. Deardorff,[‡§] Livija Medne,[‡‡] Ian D. Krantz,[‡§‡‡] Bryan L. Krock,[*†] and Avni B. Santani[*†]

From the Divisions of Genomic Diagnostics, Human Genetics,‡ and Neurology,** and the Roberts Individualized Medical Genetics Center,‡‡ Children's Hospital of Philadelphia, Philadelphia, Pennsylvania; the Departments of Pathology and Laboratory Medicine,† Pediatrics,§ and Neurology,†† Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania; the Department of Translational and Precision Medicine,¶ University of Rome Sapienza, Rome, Italy; and the Dipartimento di Medicina Molecolare e Biotecnologie Mediche,‖ Università degli Studi di Napoli Federico II, Naples, Italy*

Accepted for publication December 10, 2021.

Address correspondence to Avni B. Santani, Ph.D., Veritas Genetics, Inc., 99 Conifer Hill Dr., Danvers, MA 01923. E-mail: asantani@veritasgenetics.com.

Clinical exome sequencing (CES) aids in the diagnosis of rare genetic disorders. Herein, we report the molecular diagnostic yield and spectrum of genetic alterations contributing to disease in 700 pediatric cases analyzed at the Children's Hospital of Philadelphia. The overall diagnostic yield was 23%, with three cases having more than one molecular diagnosis and 2.6% having secondary/additional findings. A candidate gene finding was reported in another 8.4% of cases. The clinical indications with the highest diagnostic yield were neurodevelopmental disorders (including seizures), whereas immune- and oncology-related indications were negatively associated with molecular diagnosis. The rapid expansion of knowledge regarding the genome's role in human disease necessitates reanalysis of CES samples. To capture these new discoveries, a subset of cases ($n = 240$) underwent reanalysis, with an increase in diagnostic yield. We describe our experience reporting CES results in a pediatric setting, including reporting of secondary findings, reporting newly discovered genetic conditions, and revisiting negative test results. Finally, we highlight the challenges associated with implementing critical updates to the CES workflow. Although these updates are necessary, they demand an investment of time and resources from the laboratory. In summary, these data demonstrate the clinical utility of exome sequencing and reanalysis, while highlighting the critical considerations for continuous improvement of a CES test in a clinical laboratory. *(J Mol Diagn 2022, 24: 274–286; https://doi.org/10.1016/j.jmoldx.2021.12.002)*

Clinical exome sequencing (CES) is efficient and comprehensive, allowing for analysis of most protein coding regions, resulting in the diagnosis of 25% of patients referred for testing and prompting the discovery of novel disease genes.[1,2] Exome sequencing is a common molecular diagnostic test for individuals with rare genetic disorders or individuals for whom traditional diagnostic technologies were uninformative. This includes pediatric patients with clinical indications such as epilepsy, brain malformations, congenital heart disease, immunologic disorders, autism, and neurodevelopmental disabilities. CES performed as a trio (proband and parents) has the benefit of identifying *de novo* and compound heterozygous variants, while allowing

J.R.M. and A.M.I.N. contributed equally to this work.

Disclosures: A.B.S. has received licensing fees from Agilent Technologies, is a scientific advisor to the Children's Hospital of Philadelphia and 54gene, is cofounder of Opus Genomics, and holds equity in Veritas Genetics, Pathfinder, and Opus Genomics. A.M.I.N. holds equity in Veritas Genetics.

Current address of A.M.I.N and A.B.S., Veritas Genetics, Boston, MA.

https://doi.org/10.1016/j.jmoldx.2021.12.002

for the deprioritization of rare variants inherited from an unaffected parent.[1−3] CES can also identify more than one genetic condition per patient,[1,2,4,5] even when the presence of multiple conditions was not obvious based on the clinical presentation. This may include multiple pathogenic variants contributing to a patient's phenotype and secondary actionable findings.[6] Although the diagnostic yield of CES ranges from 25% to 30%, certain clinical indications are more likely to result in a diagnosis.[1−4,7−10] Moreover, facilitating timely and accurate diagnosis requires the CES laboratory to periodically perform updates and additions to existing sequencing and enrichment platforms and informatics pipelines, and to utilize new data to periodically reanalyze previously tested patients with uninformative results.[11−14] An accurate diagnosis benefits patients and their families by optimizing clinical management, predicting recurrence risks, and providing prognosis, all while ending the invasive, time-consuming, and costly diagnostic odyssey.

Herein, we provide a retrospective study of 700 pediatric patients referred to our laboratory for CES, with a variety of phenotypes and molecular diagnostic outcomes. We present a review of clinical indications prompting the test, along with their associated diagnostic rates, and discuss improvements initiated to refine and optimize the CES test. Our results demonstrate the utility of CES and reanalysis in elucidating the underlying basis of genetic disorders, while providing an assessment of CES test considerations: sequencing, informatics, troubleshooting, interpretation, and reporting.

Materials and Methods

Patient Information

This study was performed with approval from the Children's Hospital of Philadelphia Institutional Review Board. Consent for this study was not obtained because of anonymity. Between 2013 and 2017, patients were ascertained sequentially through clinical samples referred to the Division of Genomic Diagnostics Laboratory by the Roberts Individualized Medical Genetics Center, physicians and genetic counselors at Children's Hospital of Philadelphia. For each proband, informed consent was signed, which provided the option of receiving secondary findings as recommended by the American College of Medical Genetics and Genomics (ACMG).[6,15]

Clinical Indication and Scoring of Phenotypes for Statistical Correlation

At the time of testing, clinicians and genetic counselors from the Roberts Individualized Medical Genetics Center conducted a review of the clinical records and provided detailed phenotypic data, a clinical indication for testing, and a list of human phenotype ontology terms for each individual undergoing exome sequencing. Retrospectively, a genetic counselor reviewed the clinical indications and assigned 1 of 17 primary clinical indication categories: neurodevelopmental disorder, neuromuscular disorder, congenital anomalies, hearing loss, ophthalmologic issues, skeletal/connective tissue, integumentary system, respiratory disease, renal disease, liver disease, gastrointestinal disease, oncology, hematological disorder, growth disorder, suspected metabolic or mitochondrial disorder, autoimmune disease/immunodeficiency/allergies, and other (Supplemental Table S1). The other category includes various clinical indications that could not be easily incorporated into the discrete categories and does not contain shared features among the cases within this category.

In addition to assigning a primary category, the clinical indications were retrospectively used to score each individual for the presence or absence of a phenotype affecting the following body parts, organs, or systems: central nervous, musculoskeletal, integumentary, cardiovascular, immune (divided into autoimmune, immunodeficiency, or allergy), respiratory, gastrointestinal, renal, genitourinary, liver, hematological, endocrine, dental, ophthalmologic, audiological/otolaryngologic, and metabolic/biochemical. The central nervous system (CNS) and musculoskeletal systems were further subdivided to allow for more specific scoring. For example, if an individual had seizures, the individual would receive a score for both the presence of an issue impacting the CNS and the presence of seizures. Subcategories for CNS involvement included the following: nonsyndromic intellectual disability, syndromic intellectual disability, autism spectrum disorder, neurologic movement disorder, abnormality of the neural tube, abnormal brain magnetic resonance imaging, developmental delay, and seizures. Subcategories for the musculoskeletal systems included the following: neuropathy, muscular dystrophy and/or myopathy, abnormal muscle tone, skeletal, joint contractures/arthrogryposis, and connective tissue. Finally, cases were scored for the presence or absence of common reasons for referrals to a genetics specialist, such as craniofacial or dysmorphic features, growth disorders, microcephaly, macrocephaly, abnormalities noted prenatally (referred to as obstetric), structural anomalies, and malignancy (referred to as oncologic).

Extraction and Sequencing

Genomic DNA was extracted from blood following standard DNA extraction protocols in the Division of Genomic Diagnostics Laboratory at Children's Hospital of Philadelphia. A minimum of 3 μg DNA from each sample was used for whole exome library preparation and sequencing. Exome libraries were prepared using the SureSelect chemistry, following the standard manufacturer protocol (Agilent Technologies, Santa Clara, CA). Cluster generation and sequencing were performed using the TruSeq Rapid Cluster Kit-Paired-End and TruSeq Rapid SBS Kits−200 Cycle on

a HiSeq 2500 following standard manufacturer guidelines (Illumina, San Diego, CA).

Variant Calling and Quality Control

Bioinformatics Pipeline CES Version 1.0 (First 409 Exomes)

An in-house bioinformatics pipeline was developed incorporating NovoAlign (Novocraft, Selangor, Malaysia) for read alignment; Picard (Broad Institute, Cambridge, MA) for marking duplicates; and Genome Analysis Toolkit (GATK; Broad Institute, Cambridge, MA).[5,16] Best Practices for UnifiedGenotyper, with no parameter modifications,[17] was used for variant calling (reference sequence: hg19 GRCh37) and variant filtering based on read depth ($\geq$5×). Subsequent pipeline upgrades included an updated list of high-frequency pathogenic variants (managed variant list) and an internal population frequency cohort to remove sequencing artifacts.

Bioinformatics Pipeline CES Version 2.0 (Remaining 291 Exomes)

A new bioinformatics pipeline was implemented for the remaining cases (Table 1). Raw sequencing reads were aligned to the human genome reference build hg19v37 with Novoalign version 2.08.02. Duplicate reads were tagged in the alignment, and Picard HSmetrics and GATK depth of coverage were used to calculate coverage metrics for the regions of interest intervals. To help with insertion/deletion identification in the pipeline, read re-alignment base recalibration was performed around insertions/deletions using GATK version 2.2 to 5. Variant calls were generated using the GATK Unified Genotyper, and variants supported by a read depth of >5 were included in the vcf file output. Copy number variation calling was not performed.

Quality Checks

Additional quality control checks were employed to check for sample contamination, confirm the sex of the patient, and confirm the identity of the patient and any related family members when included. To rule out sample contamination during extraction and sequencing, contamination was analyzed on each sample using VerifyBamId per published methods; any BAM file found to have >2% predicted contamination was flagged and removed.[18] For contaminated samples, libraries were prepared and resequenced using a re-extracted DNA sample.

Variant Prioritization

Medical Exome

To focus the analysis on only the genes previously associated with human disease, a medically relevant gene list was developed. The medically relevant gene list is updated monthly using a custom algorithm that incorporates new gene-disease information from Online Mendelian Inheritance in Man (OMIM; *https://www.omim.org*) and the Human Genome Mutation Database (HGMD; *http://www.hgmd.cf.ac.uk/ac/index.php*) (both last accessed January 1, 2022).[19,20]

Phenoxome

To increase diagnostic efficiency, a robust, phenotype-driven model that adopts a network-based approach to facilitate automated variant prioritization, called Phenoxome, was developed.[21] This program incorporates human phenotype ontology terms into the analysis pipeline.

Variant Filtration

Variant filtering was performed using Cartagenia Bench Lab Versions 3.1 and 4.2 (Agilent Technologies), and later an in-house bioinformatics pipeline (Table 1). Included in the analysis of patient sequencing data were variants predicted to affect protein coding (missense, nonsense, frameshift, insertions, deletions, and splice site changes), along with any variants reported in HGMD. In instances where familial data were available, priority was given to variants matching an expected segregation pattern for genetic disease, which could include the following: *de novo*, homozygous, compound heterozygous, X-linked, and/or inherited segregation patterns.

Briefly, all variants in each patient's data were matched against a managed variant list containing known pathogenic variants present within the general population at high minor allele frequency (MAF), and all matching variants were retained in the analysis. Patient data were next filtered such that only variants with a <1% MAF in an internal control cohort were retained in the analysis. Next, remaining variants were grouped by their presence in HGMD

Table 1 Dynamic CES Environment: List of Bioinformatic Improvements Made Over the Course of CES Testing and Their Major Impact on Performance of the Test

Category	Updates	Impact
Variant annotation and filtration	Migration from third party to in-house software	Improved scalability
Variant calling	Software upgrades	Improved detection of large deletion and insertion variants
Minor allele frequency thresholds	Database of known high-frequency pathogenic variants	Improved variant filtration
Minor allele frequency thresholds	Expansion of internal exome cohort	Improved variant filtration
verifyBamID	Implementation of verifyBamID	Provided contamination detection

CES, clinical exome sequencing.

Professional.[19] Variants located at genomic positions matching HGMD entries were filtered such that variants with an MAF of <0.5% based on measurement of the site in >3000 chromosomes in the Exome Aggregation Consortium version 0.3.1 general population[22] or in >3000 chromosomes in the Genome Aggregation Database (gnomAD)[23] were retained in the analysis. Variants located at genomic positions not matching HGMD entries were filtered such that only variants with an MAF of <0.2% based on measurement of the site in >3000 chromosomes in the Exome Aggregation Consortium version 0.3.1 general population were retained in the analysis. All variants retained in the analysis at the conclusion of the filtration step were included in the clinical correlation step.

Variant Tiers

After filtration, each variant received a tier designation (tier 1 to 4) based on variant characteristics and/or segregation information. Tiers consisted of the following characteristics: tier 1 included variants in the probands that were i) *de novo*, ii) homozygous, iii) compound heterozygous, iv) on the X chromosome, v) loss of function, vi) listed in HGMD, vii) located in an imprinted gene, and viii) a Phenoxome gene match. Tier 2 variants included missense variants of maternal origin. Tier 3 variants were missense variants inherited from the father. Tier 4 variants were of two types: homozygous that were also homozygous in at least one parent or inherited missense with ambiguous parental origin. Variants were assigned to a single tier, with tier 1 having priority.

Exome Data Analysis

Personnel

Each exome received three levels of review after variant filtration: i) an analyst, with post-doctoral experience in genetics or molecular biology, ii) the Roberts Individualized Medical Genetics Center clinical team, and iii) a board-certified clinical molecular geneticist.

Clinical Correlation

The patient's clinical presentation, family history, and phenotype terms coded into human phenotype ontology terms are reviewed as part of the clinical correlation procedure. Clinical correlation was performed on each variant-containing gene and was conducted independently by the Roberts Individualized Medical Genetics Center and the analyst team. The correlation step focused on the assessment of the overlap between the proband's phenotype and reported disease phenotypes for that gene using OMIM and HGMD as information sources.[19,20] Each gene reviewed received one of the following designations: i) associated with phenotype; ii) possibly associated with phenotype; iii) not associated with phenotype; iv) ACMG secondary finding[6,15]; or v) candidate gene (potential human disease candidate). Variants were positively correlated (designation i or ii) if partial or significant overlap was present between the patient's phenotype and published reports on disease(s) caused by pathogenic variants in the gene. Designation iv applied only when consent for secondary findings was provided and was cocorrelated with designations i to iii. The initial analysis was restricted to variants within genes known/likely to be associated with human disease; an additional review of variants in genes with unknown medical relevance (designation v) was performed for negative cases.

Variant Classification

Variants with a clinical correlation designation other than three (not associated with phenotype) were classified for pathogenicity using an evidence-based review process centered on the ACMG standards and guidelines.[24] An analyst performed primary variant classification, which was reviewed by a board-certified clinical molecular geneticist. Variants received one of the following classifications: benign, likely benign, variant of uncertain significance (VOUS), likely pathogenic, or pathogenic.

Candidate Gene Analysis

Candidate gene analysis was included in some cases to review genes, which, at the time of analysis, were not known to cause disease in humans. Variants with either a *de novo*, compound heterozygous, homozygous, or hemizygous (male) inheritance were assessed in trios containing a proband and both unaffected parents. For trios with one affected parent, only variants inherited from the similarly affected parent were prioritized. In some instances, CES cases included multiple affected siblings; in those cases, shared variants were examined for potential disease candidacy. Exclusionary criteria for variants in genes not known to cause disease included the following: an allele frequency of >0.2% in gnomAD, presence in four or more internal cohort samples, and location in highly homologous regions. Splice variants had to be located in the ±1.2 position, and missense variants had to be highly evolutionarily conserved (through mammals and chicken). Candidate genes identified were placed into GeneMatcher.[25]

Reporting Criteria

Associated with Phenotype

This category is utilized for patients where the laboratory has identified a clear molecular diagnosis that is consistent with the clinical indication provided. The gene has been shown to cause a disease with features that almost completely overlap with the patient's phenotype. Variants that meet the following criteria are reported: pathogenic or likely pathogenic and zygosity that is consistent with the inheritance model for that disease (examples include the following: heterozygous or a *de novo* variant for an autosomal dominant condition or homozygous or compound heterozygous for an autosomal recessive disorder).

Possibly Associated with Phenotype

This category is utilized for variants in a gene associated with a disease, whose features partially overlap but may not be considered as a molecular diagnosis without additional information. Variants classified as VOUS, pathogenic, or likely pathogenic are reported. Although inheritance modeling is considered, a patient with a single heterozygous variant in a gene associated with a recessive condition is reported because the possibility of an undetected variant (such as a copy number change) on the other allele cannot be ruled out with certainty.

ACMG Secondary Findings

These findings are returned only if the patient has consented to receive ACMG secondary findings and the variants in that gene have been classified as pathogenic or likely pathogenic.

The approach to return of results for family members of a proband was implemented after discussions with the genetic counselors and ordering providers. Separate consent forms for reporting secondary findings were provided to each family member. Separate reports for the parents or other family members were not generated. If the parents gave consent for secondary findings, inheritance of the variant(s) was presented in the proband's report. Inheritance of the variant was reported as not applicable if parents did not consent to secondary findings in themselves.

Candidate Gene

For a gene to be reported as a candidate, it must not have an established association with a human genetic condition and have sufficient evidence through animal models, segregation data, functional studies, and/or biochemical pathways to suggest involvement in a disease relevant to the patient's phenotype. Various databases used to gather the appropriate information included OMIM, HGMD, Mouse Genome Informatics (MGI) Jackson Laboratory database, and PubMed. Variants were classified as variant of uncertain significance because of the limited information about gene-disease correlation.

Variant Confirmation by Sanger Sequencing

All reported variants classified as pathogenic or likely pathogenic were confirmed using Sanger sequencing before reporting (Supplemental Table S2). Variants classified as VOUS were Sanger sequenced before reporting if they were not inherited from either parent (ie, were *de novo*), were located within ± 1.2 bp of an annotated splice site, or failed to meet any of the internally validated quality criteria.

CES Reanalysis

CES reanalysis was initiated by the laboratory in 2018 on samples originally analyzed between 2014 and 2016. The CES reanalysis data set was generated by collating all genotype, phenotype, and test interpretation data generated during primary analysis for the first 300 samples that comprise the cohort. Sixty of the 300 samples had a positive diagnosis and were used for validation of the analysis. Reanalysis was performed as described previously.[11]

Statistical Analysis

For each primary clinical indication and for each phenotype subcategory, the number of patients carrying that feature with a positive outcome (reported variants that were considered strongly clinically correlated) and a negative outcome, and the number of patients without that feature with positive and negative outcome were counted. Two-sided Fisher exact tests were performed on the resulting 2×2 table to determine the significance of the association of each primary clinical indication or phenotype subcategory with a positive or negative diagnosis. The phenotype subcategories annotated for each patient were also counted, and the resulting distributions between patients with positive and negative outcomes were compared to test whether a larger number of phenotypic terms was associated with a positive outcome, using a *U*-test. In both cases, the association was defined to be significant if the test $P \leq 0.05$; multiple test correction was not applied.

The phenotype subcategories that were associated with a positive outcome (P-value threshold $= 0.05$) were then selected. All the combinations and subcombinations of these selected features were listed, and the occurrence of each of them was counted in the case of positive outcomes and overall.

Results

Primary CES analysis was performed on 700 consecutive patients. Most probands were aged <19 years (94%), with a 1.06:1 ratio of males/females. The mean age was 7.1 ± 7.6 years, with a median of 4.5 years. Traditional trios (proband and parents) made up 78% of the exomes (546/700), 4 exomes were nontraditional trios (proband, parent, and sibling), 67 exomes were duos (9.6%), 45 exomes were proband only (6.4%), 31 exomes were quad (proband, parents, and sibling), and 7 exomes were quints (proband, parents, and two siblings). A molecular diagnosis was defined as the presence of at least one pathogenic or likely pathogenic variant (diagnostic, 23%) that was clinically correlated as associated with the proband's indication for testing ($N = 162$) (Supplemental Table S2). All cases without a molecular diagnosis were considered non-diagnostic (77%), including cases with a candidate gene finding or other variants that did not meet our criteria for diagnosis. The highest diagnostic yields were seen in the 7- to 12-month (36.8%) and 18.1 to 40 years (34.0%) age groups (Figure 1). This may be explained by the 7- to 12-month group having the exome as a first-tier test or possibly this age group is starting to show or develop

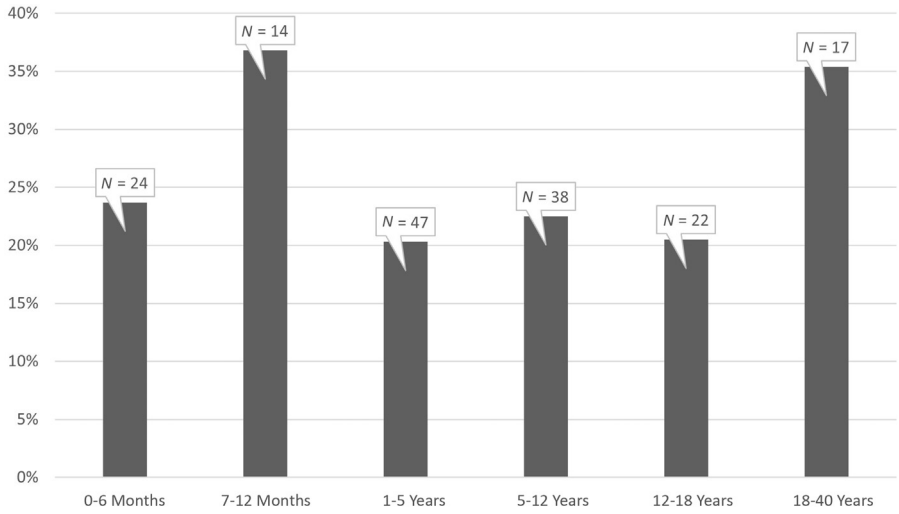

Figure 1 Diagnostic yield per age group. Ages ranged from birth to 40 years. Percentages reflect the diagnostic rate in each age group. Numbers (*N*) indicate the total number of individuals with diagnostic findings. There were no diagnostic cases aged >40 years. Diagnostic percentages for each group is as follows: 23.8% (24/101) for ages 0 to 6 months; 36.8% (14/38) for ages 7 to 12 months; 20.3% (47/231) for ages 1 to 5 years; 22.5% (38/169) for ages 5 to 12 years; 20.6% (22/107) for ages 12 to 18 years; 34.0% (17/50) for ages 18 to 40 years.

features, such as developmental delay or dysmorphic features, that could be attributed to a disorder. The 18 to 40 years age group may be explained by these probands undergoing numerous tests over the years and now exome sequencing has provided an end to the diagnostic odyssey. The diagnostic yield was 22.7% in proband-only cases, 16.4% when one additional family member was analyzed (duo), and 23.9% when two or more family members were analyzed. Of the diagnostic variants detected, 48% were *de novo* (Figure 2), and 16.8% trios had at least one diagnostic *de novo* variant reported. Among the diagnostic findings, 64% were for an autosomal dominant disorder, 24% were for autosomal recessive disorder, 11% were for X-linked disorder, and 1% were for genomic imprinting (Figure 3). Diagnostic variants included a spectrum of variant types: 39% missense and 48.9% loss of function (frameshift, large deletions, and nonsense variants) (Figure 4). Several disease genes were identified as a recurring diagnosis in three or more cases (Table 2), and more than one diagnosis was identified in three of the cases (Table 3).

Clinical indications for testing varied. Most patients were referred with a primary clinical indication related to neurodevelopmental disorders (includes seizures, 42.6%), multiple congenital anomalies (17.4%), or autoimmune disease/immunodeficiency (9.3%) (Table 4). Less common indications for referral included ophthalmologic issues (0.57%), integumentary system (0.43%), and renal disease (0.14%). Of the phenotype subcategories, CNS involvement was the most frequently observed clinical feature (63.7%), followed by intellectual disability or developmental delay (54.4%) (Table 5). Dental abnormalities were the least reported feature (0.86%). In addition, although 1.4% of patients were referred with a primary clinical indication of cancer, oncology was indicated as a feature in 2.3% of cases.

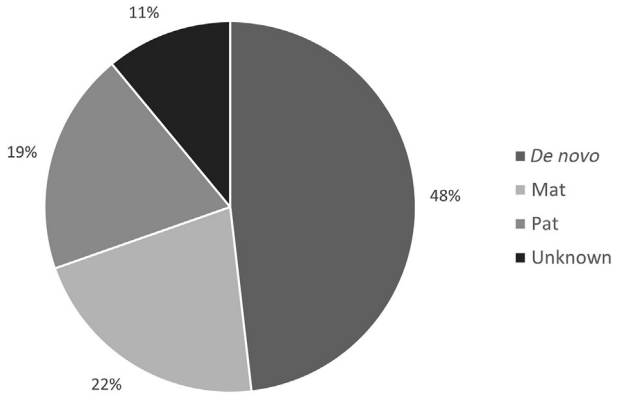

Figure 2 Parental origin of diagnostic variants. Of the 162 diagnostic cases seen in our clinical exome sequencing cohort, most (48%) were *de novo*. Maternal (Mat) and paternal (Pat) inheritance occurred at 22% and 19%, respectively. Inheritance information was unavailable for the remaining 11% of cases.

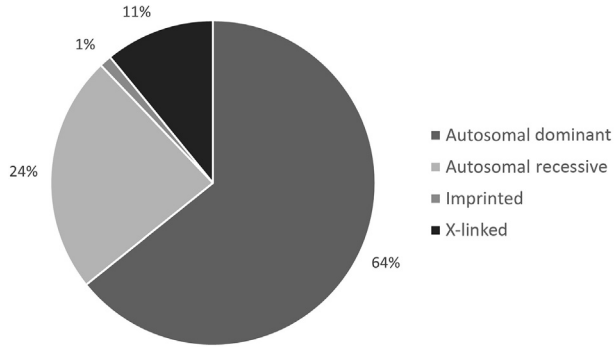

Figure 3 Inheritance patterns of molecular diagnoses. Inheritance of diagnostic variants. One percent of diagnoses were found in genes subject to genomic imprinting. $N = 162$.

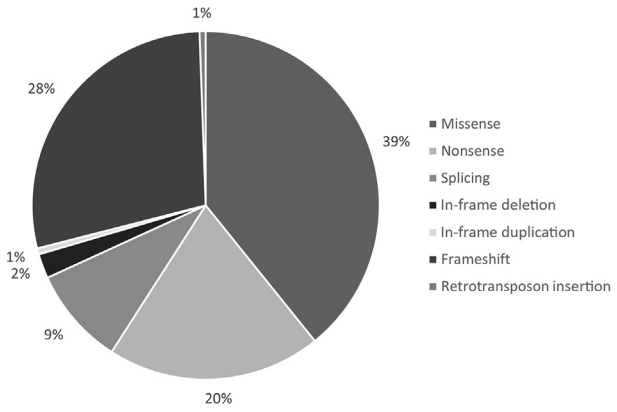

Figure 4 Type of diagnostic variants. The disease-causing variant types identified were primarily missense (39%), followed by frameshift, nonsense, and splicing. $N = 162$.

Statistical Analysis of Clinical Phenotypes

Seventeen primary clinical indication categories were compared with each other in a 2×2 Fisher exact test to determine correlation with diagnostic outcome. Neurodevelopmental disorders were correlated with a positive outcome, whereas cancer and autoimmune disease/immunodeficiency/allergies were correlated with a negative outcome (Supplemental Table S1). Similarly, comparisons were made for the 39 phenotypic subcategories to determine if any would differentiate diagnosed from undiagnosed individuals. Of the phenotypic subcategories, significant association with positive outcomes was seen in eight different categories: syndromic intellectual disability, developmental delay, CNS involvement, seizures, abnormal muscle tone, neuromuscular/musculoskeletal disorder, craniofacial involvement/dysmorphic features, and ophthalmologic findings. Significant negative associations (more often found with a negative outcome) were found with autism spectrum disorder, autoimmune disease, gastrointestinal, or oncologic findings (Supplemental Table S3). The association between the number of phenotype terms provided per patient and the frequency distributions of positive and negative diagnostic outcomes was also reviewed. In general, having a greater number of phenotype subcategories applied to a patient was predictive of a positive outcome (*U*-test $P < 0.0038$) (Figure 5).

Additional Findings

Of the 576 cases (82%) in which consent was received for reporting ACMG secondary findings,[6,15] 18 (3.1%) had reportable (pathogenic or likely pathogenic) variants (Supplemental Table S4). The most common secondary findings reported were for hereditary cancer syndromes: *APC*, *BRCA1*, *BRCA2*, *SDHB*, *TP53*, and *WT1* ($N = 7$, 38.9%), followed by cardiomyopathies ($N = 3$, 16.7%), and long-QT/Brugada syndrome ($N = 3$, 16.7%): *GLA*, *KCNQ1*, *MYL2*, *MYBPC3*, and *SCN5A*. Seven of these 18 cases (38.9%) also had a diagnostic result. Of cases with a secondary finding, 28% ($N = 5$) also had a parent with a secondary finding. The parental origin of the reported variant was not disclosed for half of the cases with a secondary finding (Supplemental Table S4). Along with secondary findings, in this cohort of 700 cases, two additional findings were reported (0.29%). These additional findings were not part of the proband's indication for CES testing, and involved compound heterozygous variants in *CFTR* and a pathogenic variant in *NR3C2*. Finally, candidate gene findings were reported in a total of 59 cases (8.4%).

CES Reanalysis

Reanalysis was performed on 240 patients from this cohort who had received nondiagnostic results from the primary CES test. Application of our reanalysis methods increased the diagnostic yield by 15.8%, which was attributed to several factors, including the following: new disease gene discovery, phenotypic expansion of known disease genes, identification of candidate genes, and updated variant classifications based on newly available literature.[11]

Discussion

To contribute to the growing body of knowledge surrounding diagnostic testing using CES, we have performed a retrospective study of 700 pediatric patients referred for CES, with correlations to outcomes for a range of phenotypes. The challenges encountered during our analysis are described below, such as secondary and additional findings, and discovery of novel disease-causing genes. We also highlight the dynamic environment, which impacts CES testing, including improvements in informatics pipelines and improved diagnostic outcomes using a cohort-based reanalysis approach.

Table 2 Recurring Diagnoses: Genetic Conditions Seen Three or More Times

Gene	Occurrences, *n*	Associated clinical syndrome
AHDC1	5	Xia-Gibbs syndrome (AD)
SYNGAP1	3	Intellectual disability (AD)
ATP1A3	4	ATP1A3-related neurologic conditions (AD)
CHD7	3	CHARGE syndrome (coloboma, *h*eart defect, choanal *a*tresia, *r*etarded growth and development, *g*enital hypoplasia, *e*ar anomalies) (AD)
EP300	3	Rubinstein-Taybi syndrome 2 (AD)
FOXG1	3	Rett syndrome, congenital variant (AD)
PMM2	3	Congenital disorder of glycosylation, type Ia (AR)

AD, autosomal dominant; AR, autosomal recessive.

Table 3 Multiple Diagnostic Findings: Three CES Cases had Two Positive Diagnostic Outcomes Related to the Clinical Indication for Testing

Sex	Age, years	Molecular Dx	Inheritance of molecular Dx	Primary clinical indication
M	16.8	*MYCN* and *TCOF1*	Unknown	Neurodevelopmental disorder*
M	6.8	*F11* and *VWF*	Maternal	Autoimmune disease/immunodeficiency/allergies
M	0.6	*AGL* and *TBX1*	Compound het and mat	Congenital anomalies

*Patient with a previously identified, *de novo*, 307-kb deletion in chromosome 22q11.21.
M, male; CES, clinical exome sequencing; Dx, diagnosis; het, heterozygous; mat, maternal.

Molecular Diagnostic Outcomes

The overall diagnostic yield for CES testing in our laboratory was 23%. Although the remaining patients were negative for a diagnostic finding, 71.3% of these negative cases still had at least one VOUS reported, and 8.4% had a candidate gene reported. Factors that yielded a higher diagnostic rate included the addition of family members to exome testing, having an autosomal dominant condition, and having a *de novo* mode of inheritance; of which the addition of parents allowed for the mode of inheritance to be determined. In addition, the presence of certain phenotype categories that yielded higher positive outcomes was also noted. In this study, primary clinical indications (main reason for CES testing) in categories of neurodevelopmental disorder were associated with a positive outcome (38.3%), whereas disorder of the immune system (13.8%) and cancer (0%) associated with a negative outcome (Supplemental Table S1).

Diagnostic findings for cases where immunodeficiency was indicated as the reason for testing included two cases of autosomal recessive *IL12RB1*-related immunodeficiency (OMIM #614891), and two cases of autosomal dominant syndromic intellectual disability, *POGZ* (OMIM #616364) and *SON* (OMIM #617140). The patient with a *de novo* *POGZ* pathogenic variant presented with combined variable immune deficiency of unknown etiology, nodular lymphoid hyperplasia of the gastrointestinal tract, hiatal hernia, gastroesophageal reflux, mild intellectual disability, hypotonia, obsessive-compulsive behaviors, polycystic ovary syndrome, and mildly dysmorphic features. The patient with a *SON* pathogenic variant presented with hearing loss, developmental delay, failure to thrive, intermittent diarrhea, acute pyelonephritis, and presumed immunodeficiency. Although both cases had a primary clinical indication of immune disorder, the presence of additional phenotypic terms at the time of test requisition contributed to the molecular diagnosis. This is in agreement with our observation that more phenotype terms are predictive of a diagnostic outcome (Figure 5). Although the overall diagnostic rate for immune disorders is presently low, they still comprised 10% of tests ordered, indicating a need to obtain molecular diagnoses in this population. This type of cohort may benefit from reanalysis as novel gene disease associations are discovered, which may improve diagnostic rates over time.

Table 4 Frequency of Primary Clinical Indications for CES Testing

Primary clinical indication	Value, *n* (%)
Neurodevelopmental disorder	298 (42.6)
Congenital anomalies	122 (17.4)
Autoimmune disease/immunodeficiency/allergies	65 (9.3)
Other	53 (7.6)
Suspected metabolic or mitochondrial disorder	27 (3.9)
Neuromuscular disorder	27 (3.9)
Gastrointestinal disease	21 (3.0)
Skeletal/connective tissue	18 (2.6)
Hearing loss	14 (2.0)
Hematological disorder	11 (1.6)
Cancer	10 (1.4)
Respiratory disease	10 (1.4)
Growth disorder	10 (1.4)
Liver disease	6 (0.9)
Ophthalmologic issues	4 (0.6)
Integumentary system	3 (0.4)
Renal disease	1 (0.1)

Primary clinical indications were itemized into 17 categories. Patients were assigned to one primary category based on the main indication for exome testing.
CES, clinical exome sequencing.

The outcomes for phenotype subcategories were studied next (Supplemental Table S3). Syndromic intellectual disability, abnormal muscle tone, craniofacial dysmorphism, developmental delay, CNS involvement, neuromuscular/musculoskeletal, ophthalmologic issues, and seizures were all associated with a positive outcome, whereas autoimmunity, gastrointestinal, autism spectrum disorder, and oncology were associated with a negative outcome. In addition to reviewing individual primary clinical indications and phenotype subcategories, their combinations were studied and the results were as expected; terms associated with positive outcomes were positive when combined, and terms associated with negative outcomes were negative when combined (data not shown). Interestingly, there was no significant association with outcome when positively and negatively associated terms were combined.

The positive outcome rates for these clinical features (Supplemental Tables S2 and S3) are comparable to previous CES cohort studies,[1,2,7,26] indicating consistency in detection and reporting of their known genetic causes. In addition, this study added statistical analysis to see if they can explain observed outcomes. Although these findings indicate the clinical indications that are more or less likely to yield a diagnosis, there are a couple of caveats. For instance, some of these phenotype categories have a small number of

Table 5 Frequency of Phenotype Subcategories for CES Testing

Phenotype subcategory	Value, *n* (%)
CNS involvement	446 (63.7)
Developmental delay	329 (54.4)
Neuromuscular/musculoskeletal	322 (46.0)
Structural birth defect(s)	236 (33.7)
Growth abnormality	210 (30.0)
Craniofacial involvement/dysmorphic features	209 (29.9)
Abnormal muscle tone	172 (24.6)
Gastrointestinal	155 (22.1)
Skeletal	123 (17.6)
Seizures	119 (17.0)
Abnormal brain MRI	115 (16.4)
Cardiovascular	108 (15.4)
Ophthalmologic	104 (14.9)
Audiologic/otolaryngolic	83 (11.9)
Respiratory/pulmonary	82 (11.7)
Microcephaly	80 (11.4)
Immunodeficiency (rare and/or recurrent infections)	78 (11.1)
Neurologic movement disorder (ataxia/spasticity, tremor, dystonia, parkinsonism, or myoclonus)	69 (9.9)
Autism spectrum disorder	66 (9.4)
Endocrine	63 (9.0)
Connective tissue	59 (8.4)
Hematological/vascular	58 (8.3)
Genitourinary	49 (7.0)
Renal	47 (6.7)
Integumentary (includes hair, skin, and nails)	42 (6.0)
Syndromic intellectual disability	42 (6.0)
Metabolic/biochemical	39 (5.6)
Macrocephaly	35 (5.0)
Autoimmune disease	34 (4.9)
Muscular dystrophy and/or myopathy	32 (4.6)
Liver	31 (4.4)
Joint contracture/arthrogryposis multiplex congenita	23 (3.3)
Oncologic	16 (2.3)
Neuropathy	15 (2.1)
Allergies	15 (2.1)
Abnormality of spinal cord/neural tube	13 (1.9)
Nonsyndromic intellectual disability	11 (1.6)
Obstetric	11 (1.6)
Dental	6 (0.9)

Individual clinical features were listed in 39 subcategories. Each patient was assigned to one or more subcategories.

CES, clinical exome sequencing; CNS, central nervous system; MRI, magnetic resonance imaging.

cases and results may be inconclusive. Therefore, it would be useful to apply this analysis to an even larger cohort and obtain statistically significant results across a broader range of phenotypic terms. Furthermore, it is possible that the phenotypes associated with negative outcomes do not have extensive gene-disease etiologies in the literature. These negatively associated clinical indications are likely to be candidates for positive reanalysis findings, as more information becomes available regarding their gene-disease associations.

A common challenge that laboratories face while analyzing CES cases, with a wide phenotypic spectrum, is ensuring that the analysis strategies identify multiple molecular etiologies, which is observed in approximately 1% to 4% of CES cases.[1–4,7] In some cases, multiple overlapping features were present between both genetic diagnoses, making it challenging to delineate the exact molecular etiology of each clinical feature. An example of one such case is an individual with seizures, hypotonia, microcephaly, colobomas, proptosis, optic nerve hypoplasia, dysmorphic features, and intellectual disability who carried heterozygous variants in both *MYCN* (OMIM #164840) and *TCOF1* (OMIM #606847) along with a previously identified *de novo* 307-kb deletion of chromosome 22q11.21, including only the low copy region B.[27]

Comprehensive analysis is necessary to ensure that laboratories do not stop prematurely before all diagnoses are detected. In such cases, laboratories must continue the analysis of additional phenotypically relevant variants to rule out the possibility of an additional molecular diagnosis. In a case that highlights this challenge, the proband presented with complex brain malformation and abnormal electroencephalography with seizures. A *de novo* missense variant was identified in *GRIN2B* (OMIM #138252). At the time of analysis, the published literature reported patients with *GRIN2B* variants and seizures only, leading to the question of whether the brain malformations in our patient represented an expansion of the *GRIN2B* phenotype, or if there was a second, unknown genetic diagnosis with possible recurrence risk ramifications for the family. Since then, a *GRIN2B* cohort demonstrating cortical malformations has been published.[28] This case highlights the complexity of a clinical laboratory's need to balance turnaround time while maintaining a sensitive and comprehensive approach.

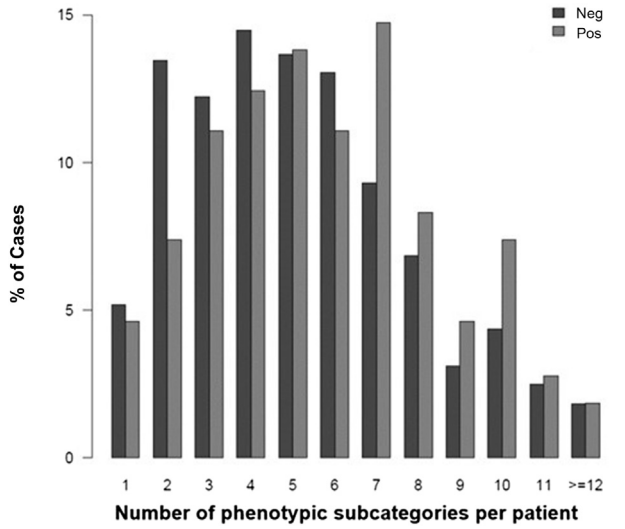

Figure 5 Outcome per number of phenotypic terms. Percentage of positive (Pos) and negative (Neg) outcomes per number of phenotype subcategories.

Another finding of interest involves the inclusion of clinical information from family members and its utility in testing. As expected, trio testing was beneficial for capturing segregation information and to add weight to a variant's pathogenicity, especially with novel missense changes that might otherwise be classified as VOUS and possibly considered nondiagnostic. Trio-based exome analysis also helped deprioritize heterozygous variants from a seemingly unaffected parent. However, a maternally inherited, pathogenic variant was encountered in *FOXG1* (OMIM #164874). The proband presented with microcephaly, developmental delay, hypotonia, failure to thrive, delayed myelination, complex partial seizures, and structural brain abnormalities. The mother was unaffected, and although she carried the variant, she was found to be mosaic, a state that has repercussions for recurrence risk and need for further genetic counseling.

Overall, these findings demonstrate the need for a comprehensive CES approach that can anticipate and adequately address a variety of diagnostic challenges, including diverse and complex clinical indications, mosaicism, and multiple findings.

Complexities with Reporting Secondary and Additional Findings

When the ACMG recommendations for reporting of secondary findings were first published,[6] there were concerns regarding appropriate procedures for consenting of patients and family members.[29] Parents may provide consent for their child to undergo secondary findings analysis; however, if a secondary finding is identified in the child, then it is likely to have been inherited. Therefore, parents analyzed as part of a trio would put the laboratory in the position of seeing data that will have implications not only for the patient, but also for the parents. For adults, additional considerations include making arrangements should they choose to obtain additional health, life, and/or other supplemental insurance before release of results. This issue led us to reevaluate our consent process such that parents would have the option to consent to receive their own status for secondary findings identified in their child.

The delivery mechanism for return of results relevant to the parent's health was also evaluated. Parental inheritance may be included in the proband's report (with appropriate consent); however, a separate report for the parents would ensure that this information could be stored in their individual electronic health records, thereby promoting appropriate clinical management. Although facilitating consent and multiple return of results is an ideal scenario for the patient and family members, it is important to recognize the impact on the laboratory. A case management workflow was instituted for this scenario (including separate accessioning, tracking of consent options, and confirmation of findings by an orthogonal technique, and a separate reporting workflow). Because the percentage of patients who receive an actionable finding is about 1.7% to 6.2%,[1,30] it is essential to ameliorate the impact on the laboratory, as well as the clinicians, with appropriate workflow management tools, while still providing the best care for patients and their families.

In this cohort, compound heterozygous pathogenic variants were identified in *CFTR* (OMIM #602421), along with a pathogenic variant in *NR3C2* (OMIM #600983) associated with pseudohypoaldosteronism (OMIM #177735) (Supplemental Table S4). The proband with the pathogenic variant in *NR3C2* had moderately short stature, sacral agenesis, failure to thrive, hypotonia, and hypertelorism listed as the clinical indication for exome testing. The proband medical record was examined in greater detail, and it was thought that this variant provided an explanation for the proband's salt wasting, a feature that was part of the proband's medical history, but was not listed as an indication for CES. The *CFTR* variants were found in a patient with hypogammaglobulinemia, lymphopenia, primary intestinal lymphangiectasia, and lower limb asymmetry as the clinical indication for exome sequencing. The proband was aged 5 years at testing and had passed all newborn screening. These *CFTR* variants are not associated with typical cystic fibrosis but associated with an increased risk for pancreatitis. Therefore, it is important to be aware that additional findings might be part of a patient's medical history and might provide changes to medical management, but may not be included as part of the indication for testing.[31] Therefore, appropriate consent strategies are important to educate the patients before test initiation.

Complexities with Reporting and Publishing Novel Candidate Gene Findings

Candidate gene findings were reported in 8.4% of cases. Reporting of candidate genes from whole exome sequencing has been debated among clinical laboratories, particularly because candidate genes are not conclusively proven to cause disease[32] and the evidence needed to assess clinical validity of a gene to a disease in the form of peer-reviewed publications tends to be limited. The risks associated with reporting candidate genes in CES may include cessation of further genetic testing, initiation of prenatal testing, and other medical management decisions based on this limited evidence. On the other hand, using the power of CES can help to elucidate unique and novel molecular diagnoses. For example, of the 59 cases that received candidate gene findings, 7 (12.9%) subsequently received updated reports where the candidate was upgraded to a disease-associated gene.[16,33–37]

For these reasons, a standardized system for reporting potential candidate genes was developed. This policy defines candidate genes as those without current association to human disease, but for which significant experimental evidence, generated using either *in vitro* or *in vivo* models, exists, suggesting a possible role in human disease. Given the limited published evidence supporting the role of these

genes in human disease, all variants in candidate genes are classified and reported as VOUS. Candidate genes are reported in a separate table in the final report with a clear stipulation that medical decisions should not be made based on this limited evidence. The report clearly states that additional case reports, segregation studies, and/or functional data are required to determine the significance of the candidate findings within the context of human disease.

This approach has allowed for several international collaborations, resulting in publications, and six of our previously reported candidate genes are now associated with disease. Some of these candidate to positive exomes were a result of a positive match in GeneMatcher.[25,38] In addition, some of the candidate genes were of interest to researchers and clinicians here at the Children's Hospital of Philadelphia and have resulted in several publications linking these genes to disease.[16,34]

Despite these successes, the gene discovery process is labor intensive, and requires consistent follow-up and communication. An organized infrastructure is necessary to manage the multiple gene submissions to GeneMatcher, the corresponding phenotypes, the potential multi-institution matches, and tracking of the numerous communications between various groups. In this situation, a clinical laboratory, because of its central role in gene identification and access to additional patients, may be a logical facilitator of these discussions between clinicians and research collaborators. This undertaking is difficult to measure in terms of time and resources and is likely to increase with the number of exomes performed over time. Our experience suggests that taking the initial gene discovery to a publication requires a team effort with clinicians, researchers, and clinical laboratories; and transparent and effective communication is essential to ensure a successful and productive collaboration.

CES Reanalysis

The rapid pace at which new disease genes and disease-causing variants are identified supports the idea that conducting periodic reanalysis of nondiagnostic CES samples is likely to reveal novel molecular diagnoses.[12,13] As part of developing a semiautomated approach to provide reanalysis for CES patients, a cohort of 240 nondiagnostic cases were selected by the laboratory for reanalysis. To this end, we previously reported the application of a reanalysis method that increased the diagnostic yield of a previously negative cohort.[11] Although the rationale for reanalysis of nondiagnostic CES samples is compelling from the standpoint of improving patient care, there are significant barriers to implementing reanalysis procedures that can be scaled with an expanding patient cohort while simultaneously minimizing impacts on laboratory turnaround time and labor. Reanalysis can include all or some of several modalities, such as the following: i) utilization of updated patient phenotype information and additional family members, ii) improved informatics algorithms (eg, copy number calling from short read sequence data), and/or iii) updated exome capture and resequencing. Although estimates of diagnostic yield and the workload required for reanalysis will likely vary between clinical laboratories and cohorts, defining the scope of reanalysis can aid clinical laboratories in determining the frequency and resources with which to reanalyze their data. Our experience suggests that for reanalysis to be successful, clinical laboratories should carefully consider the resources required for the implementation of a reanalysis infrastructure, including adequate personnel and informatics support, education of ordering clinicians, and development of strategies to facilitate timely reimbursement of these types of tests.

Keeping Up with Advancements in Genomics in a Dynamic Environment

One striking feature of CES at our site has been its dynamic nature, driven primarily by the desire to enhance this clinical test with improved methods, while balancing high clinical sensitivity and operational scalability. An example of this is the evolution of pipelines and tools used to perform CES in this laboratory (Table 1). At the initial test launch, third-party software was utilized for variant annotation and filtration. Increasing test volumes to improve efficiency and reduce costs led to the development and validation of an in-house pipeline. Another approach to improving efficiency was to improve variant filtration strategies by refining the MAF threshold by developing a genome-wide knowledge-base of high-frequency pathogenic variants and by building a database of internal exomes. Additional modifications were aimed at improving test quality, such as upgrading from GATK unified genotype to GATK HaplotypeCaller[39] in CWES2.0, which resulted in improved calling of larger insertion and deletion variants, and implementation of verifyBamID to prospectively identify sample contamination.[18] It is clear that in a CES laboratory, processes will continue to evolve over time, although frequent validation of new processes and new tools presents a logistical challenge for a hospital-based diagnostic laboratory. It is nonetheless an essential task to maintain the highest quality test possible.

In summary, CES is an unbiased test with proven clinical utility to diagnose patients with severe and/or multisystemic conditions. Guiding ordering clinicians by suggesting helpful human phenotype ontology terms and revealing the most promising clinical indications will increase diagnostic yield, make the diagnostic process more efficient, and reduce costs. In addition, the laboratory has a responsibility to provide an accurate and efficient test while ensuring that the processes and workflows are up-to-date.

Supplemental Data

Supplemental material for this article can be found at *http://doi.org/10.1016/j.jmoldx.2021.12.002*.

References

1. Retterer K, Juusola J, Cho MT, Vitazka P, Millan F, Gibellini F, Vertino-Bell A, Smaoui N, Neidich J, Monaghan KG, McKnight D, Bai R, Suchy S, Friedman B, Tahiliani J, Pineda-Alvarez D, Richard G, Brandt T, Haverfield E, Chung WK, Bale S: Clinical application of whole-exome sequencing across clinical indications. Genet Med 2016, 18:696−704
2. Farwell KD, Shahmirzadi L, El-Khechen D, Powis Z, Chao EC, Tippin Davis B, Baxter RM, Zeng W, Mroske C, Parra MC, Gandomi SK, Lu I, Li X, Lu H, Lu HM, Salvador D, Ruble D, Lao M, Fischbach S, Wen J, Lee S, Elliott A, Dunlop CL, Tang S: Enhanced utility of family-centered diagnostic exome sequencing with inheritance model-based analysis: results from 500 unselected families with undiagnosed genetic conditions. Genet Med 2015, 17: 578−586
3. Lee H, Deignan JL, Dorrani N, Strom SP, Kantarci S, Quintero-Rivera F, Das K, Toy T, Harry B, Yourshaw M, Fox M, Fogel BL, Martinez-Agosto JA, Wong DA, Chang VY, Shieh PB, Palmer CG, Dipple KM, Grody WW, Vilain E, Nelson SF: Clinical exome sequencing for genetic identification of rare Mendelian disorders. JAMA 2014, 312:1880−1887
4. Yang Y, Muzny DM, Reid JG, Bainbridge MN, Willis A, Ward PA, Braxton A, Beuten J, Xia F, Niu Z, Hardison M, Person R, Bekheirnia MR, Leduc MS, Kirby A, Pham P, Scull J, Wang M, Ding Y, Plon SE, Lupski JR, Beaudet AL, Gibbs RA, Eng CM: Clinical whole-exome sequencing for the diagnosis of mendelian disorders. N Engl J Med 2013, 369:1502−1511
5. Lazaridis KN, Schahl KA, Cousin MA, Babovic-Vuksanovic D, Riegert-Johnson DL, Gavrilova RH, McAllister TM, Lindor NM, Abraham RS, Ackerman MJ, Pichurin PN, Deyle DR, Gavrilov DK, Hand JL, Klee EW, Stephens MC, Wick MJ, Atkinson EJ, Linden DR, Ferber MJ, Wieben ED, Farrugia G; Individualized Medicine Clinic Members: Outcome of whole exome sequencing for diagnostic odyssey cases of an individualized medicine clinic: the Mayo Clinic experience. Mayo Clin Proc 2016, 91:297−307
6. Green RC, Berg JS, Grody WW, Kalia SS, Korf BR, Martin CL, McGuire AL, Nussbaum RL, O'Daniel JM, Ormond KE, Rehm HL, Watson MS, Williams MS, Biesecker LG; American College of Medical Genetics and Genomics: ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. Genet Med 2013, 15:565−574
7. Trujillano D, Bertoli-Avella AM, Kumar Kandaswamy K, Weiss ME, Koster J, Marais A, Paknia O, Schroder R, Garcia-Aznar JM, Werber M, Brandau O, Calvo Del Castillo M, Baldi C, Wessel K, Kishore S, Nahavandi N, Eyaid W, Al Rifai MT, Al-Rumayyan A, Al-Twaijri W, Alothaim A, Alhashem A, Al-Sannaa N, Al-Balwi M, Alfadhel M, Rolfs A, Abou Jamra R: Clinical exome sequencing: results from 2819 samples reflecting 1000 families. Eur J Hum Genet 2017, 25:176−182
8. Tarailo-Graovac MSC, Ross CJ, Horvath GA, Salvarinova R, Ye XC, Zhang LH, et al: Exome sequencing and the management of neurometabolic disorders. N Engl J Med 2016, 374:2246−2255
9. Powis ZFHK, Speare V, Cain T, Blanco K, Mowlavi LS, Mayerhofer EM, Tilstra D, Vedder T, Hunter JM, Tsang M, Gonzalez L, Vockley G, Tang S: Exome sequencing in neonates: diagnostic rates, characteristics, and time to diagnosis. Genet Med 2018, 20:1468−1471
10. Meng LPM, Saronwala A, Magoulas P, Ghazi AR, Vetrini F, Zhang J, et al: Use of exome sequencing for infants in intensive care units: ascertainment of severe single-gene disorders and effect on medical management. JAMA Pediatr 2017, 171:e173438
11. Baker SW, Murrell JR, Nesbitt AI, Pechter KB, Balciuniene J, Zhao X, Yu Z, Denenberg EH, DeChene ET, Wilkens AB, Bhoj EJ, Guan Q, Dulik MC, Conlin LK, Abou Tayoun AN, Luo M, Wu C, Cao K, Sarmady M, Bedoukian EC, Tarpinian J, Medne L, Skraban CM, Deardorff MA, Krantz ID, Krock BL, Santani AB: Automated clinical exome reanalysis reveals novel diagnoses. J Mol Diagn 2019, 21:38−48
12. Ewans LJSD, Shrestha R, Zhu Y, Gayevskiy V, Ying K, Walsh C, Lee E, Kirk EP, Colley A, Ellaway C, Turner A, Mowat D, Worgan L, Freckmann ML, Lipke M, Sachdev R, Miller D, Field M, Dinger ME, Buckley MF, Cowley MJ, Roscioli T: Whole-exome sequencing reanalysis at 12 months boosts diagnosis and is cost-effective when applied early in Mendelian disorders. Genet Med 2018, 20:1564−1574
13. Wenger AMGH, Bernstein JA, Bejerano G: Systematic reanalysis of clinical exome data yields additional diagnoses: implications for providers. Genet Med 2017, 19:209−214
14. Roy SCC, Karunamurthy A, Kip NS, Klee EW, Lincoln SE, Leon A, Pullambhatla M, Temple-Smolkin RL, Voelkerding KV, Wang C, Carter AB: Standards and guidelines for validating next-generation sequencing bioinformatics pipelines: a joint recommendation of the Association for Molecular Pathology and the College of American Pathologists. J Mol Diagn 2018, 20:4−27
15. Kalia SS, Adelman K, Bale SJ, Chung WK, Eng C, Evans JP, Herman GE, Hufnagel SB, Klein TE, Korf BR, McKelvey KD, Ormond KE, Richards CS, Vlangos CN, Watson M, Martin CL, Miller DT: Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American College of Medical Genetics and Genomics. Genet Med 2017, 19:249−255
16. Skraban CM, Wells CF, Markose P, Cho MT, Nesbitt AI, Au PYB, et al: WDR26 haploinsufficiency causes a recognizable syndrome of intellectual disability, seizures, abnormal gait, and distinctive facial features. Am J Hum Genet 2017, 101:139−148
17. Gibson KMNA, Cao K, Yu Z, Denenberg E, DeChene E, Guan Q, Bhoj E, Zhou X, Zhang B, Wu C, Dubbs H, Wilkens A, Medne L, Bedoukian E, White PS, Pennington J, Luo M, Conlin L, Monos D, Sarmady M, Marsh E, Zackai E, Spinner N, Krantz I, Deardorff M, Santani A: Novel findings with reassessment of exome data: implications for validation testing and interpretation of genomic data. Genet Med 2018, 20:329−336
18. Jun GFM, Hetrick KN, Romm JM, Doheny KF, Abecasis GR, Boehnke M, Kang HM: Detecting and estimating contamination of human DNA samples in sequencing and array-based genotype data. Am J Hum Genet 2012, 91:839−848
19. Stenson PD, Ball EV, Mort M, Phillips AD, Shiel JA, Thomas NS, Abeysinghe S, Krawczak M, Cooper DN: Human Gene Mutation Database (HGMD): 2003 update. Hum Mutat 2003, 21:577−581
20. Amberger J, Bocchini CA, Scott AF, Hamosh A: McKusick's Online Mendelian Inheritance in Man (OMIM). Nucleic Acids Res 2009, 37: D793−D796
21. Wu C, Devkota B, Evans P, Zhao X, Baker SW, Niazi R, Cao K, Gonzalez MA, Jayaraman P, Conlin LK, Krock BL, Deardorff MA, Spinner NB, Krantz ID, Santani AB, Tayoun ANA, Sarmady M: Rapid and accurate interpretation of clinical exomes using Phenoxome: a computational phenotype-driven approach. Eur J Hum Genet 2019, 27:612−620
22. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al; Exome Aggregation Consortium: Analysis of protein-coding genetic variation in 60,706 humans. Nature 2016, 536: 285−291
23. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al: The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 2020, 581:434−443
24. Richards SAN, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL; ACMG Laboratory Quality Assurance Committee: 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, 17:405−424

25. Sobreira N, Schiettecatte F, Valle D, Hamosh A: GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. Hum Mutat 2015, 36:928−930
26. Arts P, Simons A, AlZahrani MS, Yilmaz E, AlIdrissi E, van Aerde KJ, et al: Exome sequencing in routine diagnostics: a generic test for 254 patients with primary immunodeficiencies. Genome Med 2019, 11
27. McDonald-McGinn DM, Hain HS, Emanuel BS, Zackai EH: 22q11.2 Deletion syndrome. In GeneReviews [Internet]. Copyright University of Washington, Seattle. 2020. Available at: https://www.ncbi.nlm.nih.gov/books/NBK1523 (last revised February 27, 2020)
28. Platzer KYH, Schütz H, Winschel A, Chen W, Hu C, Kusumoto H, et al: GRIN2B encephalopathy: novel findings on phenotype, variant clustering, functional consequences and treatment aspects. J Med Genet 2017, 54:460−470
29. Roche MI, Berg JS: Incidental findings with genomic testing: implications for genetic counseling practice. Curr Genet Med Rep 2015, 3:166−176
30. Hart MRBB, Blout CL, Christensen KD, Amendola LM, Bergstrom KL, Biswas S, et al: Secondary findings from clinical genomic sequencing: prevalence, patient perspectives, family history assessment, and health-care costs from a multisite study. Genet Med 2019, 21:1100−1110
31. Saelaert MMH, Moerenhout T, De Baere E, Devisch I: Criteria for reporting incidental findings in clinical exome sequencing - a focus group study on professional practices and perspectives in Belgian genetic centres. BMC Med Genomics 2019, 12:123
32. Iglesias AA-YK, Wynn J, Wilson A, Truitt Cho M, Guzman E, Sisson R, Egan C, Chung WK: The usefulness of whole-exome sequencing in routine clinical practice. Genet Med 2014, 16:922−931
33. Knierim EGE, Seifert F, Morales-Gonzalez S, Unudurthi SD, Hund TJ, Stenzel W, Schuelke M: A recessive mutation in beta-IV-spectrin (SPTBN4) associates with congenital myopathy, neuropathy, and central deafness. Hum Genet 2017, 136:903−910
34. Küry S, van Woerden GM, Besnard T, Proietti Onori M, Latypova X, Towne MC, et al: De novo mutations in protein kinase genes CAMK2A and CAMK2B cause intellectual disability. Am J Hum Genet 2017, 101:768−788
35. Meyer ECK, Rankin J, Nichols JM, Grozeva D, Joseph AP, Mencacci NE, et al: Mutations in the histone methyltransferase gene KMT2B cause complex early-onset dystonia. Nat Genet 2017, 49:223−237
36. Zech MBS, Maier EM, Borggraefe I, Vill K, Laccone F, Pilshofer V, Ceballos-Baumann A, Alhaddad B, Berutti R, Poewe W, Haack TB, Haslinger B, Strom TM, Winkelmann J: Haploinsufficiency of KMT2B, encoding the lysine-specific histone methyltransferase 2B, results in early-onset generalized dystonia. Am J Hum Genet 2016, 99:1377−1387
37. Dias CES, Graham SA, McRae J, Sawiak SJ, Hurst JA, Joss SK, Holder SE, Morton JE, Turner C, Thevenon J, Mellul K, Sánchez-Andrade G, Ibarra-Soria X, Deriziotis P, Santos RF, Lee SC, Faivre L, Kleefstra T, Liu P, Hurles ME; DDD Study, Fisher SE, Logan DW: BCL11A haploinsufficiency causes an intellectual disability syndrome and dysregulates transcription. Am J Hum Genet 2016, 99:253−274
38. Olson HE, Jean-Marcais N, Yang E, Heron D, Tatton-Brown K, van der Zwaag PA, et al: A recurrent de novo PACS2 heterozygous missense variant causes neonatal-onset developmental epileptic encephalopathy, facial dysmorphism, and cerebellar dysgenesis. Am J Hum Genet 2018, 103:631
39. Ryan Poplin VR-R, DePristo MA, Fennell TJ, Carneiro MO, Van der Auwera GA, Kling DE, Gauthier LD, Levy-Moonshine A, Roazen D, Shakir K, Thibault J, Chandran S, Whelan C, Lek M, Gabriel S, Daly MJ, Neale B, MacArthur DG, Banks E: Scaling accurate genetic variant discovery to tens of thousands of samples. bioRxiv 2018, [Preprint], doi:10.1101/201178