Full Title: Genetic, demographic and clinical variables act synergistically to impact neurodevelopmental outcomes in children with single ventricle heart disease

Authors:

Thomas A. Miller*, Department of Pediatrics, Maine Medical Center, Portland, ME; Department of Pediatrics, University of Utah, Salt Lake City, UT

Edgar J. Hernandez*, Department of Human Genetics and Utah Center for Genetic Discovery, University of Utah, Salt Lake City, UT

J. William Gaynor, Department of Surgery, Children's Hospital of Philadelphia, and the Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA

Mark W. Russell, Department of Pediatrics, University of Michigan, MI

Jane W. Newburger, MD, MPH, Department of Cardiology, Boston Children's Hospital, Department of Pediatrics, Harvard Medical School, Boston, MA

Wendy Chung, Departments of Pediatrics and Medicine, Columbia University, New York, NY

Elizabeth Goldmuntz, Division of Cardiology, Children's Hospital of Philadelphia, Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA

James F. Cnota, Heart Institute, Cincinnati Children's Hospital, Cincinnati, OH

Sinai C. Zyblewski, Department of Pediatrics, Medical University of South Carolina, Charleston, SC

William T. Mahle, Children's Healthcare of Atlanta, Atlanta, GA

Victor Zak, Healthcore, Watertown, MA

Chitra Ravishankar, Department of Pediatrics, Children's Hospital of Philadelphia, Philadelphia, PA.

Jonathan R. Kaltman, Division of Cardiology, Children's National Hospital, Washington, DC

Brian W. McCrindle, Labatt Family Heart Centre, The Hospital for Sick Children, Department of Pediatrics, University of Toronto, Toronto, ON

Shanelle Clarke, Department of Pediatrics Emory University School of Medicine, Atlanta Georgia

Jodie K. Votava-Smith, Department of Pediatrics, Children's Hospital Los Angeles, Los Angeles, CA

Eric M. Graham, MD, Department of Pediatrics, Medical University of South Carolina, Charleston, SC

Mike Seed, MD, Division of Cardiology, Department of Paediatrics, University of Toronto and Hospital for Sick Children, Toronto, ON, Canada

Nancy Rudd, Department of Pediatrics, Medical College of Wisconsin, Milwaukee, WI

Daniel Bernstein, M.D. Stanford University School of Medicine, Stanford, CA.

Teresa M. Lee, Department of Pediatrics, Columbia University, New York, NY

Mark Yandell, Department of Human Genetics and Utah Center for Genetic Discovery, University of Utah, Salt Lake City, UT

Martin Tristani-Firouzi, Department of Pediatrics, University of Utah, Salt Lake City, UT

*These authors contributed equally

Email for correspondence:

Thomas.A.Miller@mainehealth.org

myandell@genetics.utah.edu

Martin.Tristani@utah.edu

The content of this article is the responsibility of the authors and do not represent the official views of the National Heart, Lung, and Blood Institute or the NIH.

ABSTRACT:

Recent large-scale sequencing efforts have shed light on the genetic contribution to the etiology of congenital heart defects (CHD); however, the relative impact of genetics on clinical outcomes remains less understood. Outcomes analyses using genetic data are complicated by the intrinsic severity of the CHD lesion and by interactions with conditionally dependent clinical variables. Here we apply Bayesian Networks, an explainable Artificial Intelligence solution, to describe the intertwined relationships between clinical variables, demography, and genetics in a cohort of children with single ventricle CHD. As isolated variables, a damaging genetic variant in a gene related to abnormal heart morphology and prolonged ventilator support following stage I palliative surgery increased the probability of having a low Mental Developmental Index (MDI) score at 14 months of age by 1.9- and 5.8-fold, respectively. However, in combination, these variables acted synergistically to further increase the probability of a low MDI score by 10-fold. Likewise, genetic information was predictive of a favorable neurodevelopmental outcome. For example, the absence of a damaging variant in a known syndromic CHD gene and a shorter post-operative ventilator support increased the probability of a normal MDI score 1.7- and 2.4-fold, respectively, but in combination increased the probability of a good outcome by 59-fold. Our analyses suggest a modest genetic contribution to neurodevelopmental and growth outcomes as isolated variables, similar to known clinical predictors. By contrast, genetic, demographic, and clinical variables interact synergistically to markedly impact clinical outcomes. These findings underscore the importance of capturing and quantifying the impact of damaging genomic

variants in the context of multiple, conditionally dependent variables, such as pre- and post-operative factors, and demography.

INTRODUCTION

While survival for infants with single ventricle congenital heart disease (SV-CHD) has improved dramatically, significant healthcare utilization, morbidity, and mortality persist.[1] Despite improvements in standardized care, neurodevelopmental and growth outcomes vary substantially in children with SV-CHD.[2] Identifiable syndromes and chromosomal abnormalities are associated with worse outcomes for multiple forms of complex CHD.[3-6] Predicted damaging genetic variants in genes related to CHD may contribute to poor outcomes, irrespective of whether the patient has syndromic CHD. Outside of aneuploidy and other known syndromes, the genetic architecture of CHD includes high-impact, rare variants across many loci. Thus, many variants that contribute to CHD have a small population attributable risk (PAR).[7,8] This low PAR makes determination of the association of variants with outcomes difficult, especially when studying relatively small numbers of individuals with SV-CHD. As a result, the relative contribution of genomic variation to neurodevelopment and growth outcomes remains poorly described, especially in the context of clinical variables and other comorbidities, which can combine with one another in a conditionally dependent manner to create a constellation of influence on any given outcome.[9] Two clinical variables are conditionally dependent if the probability of encountering both variables is not equal to the multiplication of their individual probabilities. For example, sex and bicuspid aortic valve (BAV) are

conditionally-dependent, in that the probability of encountering an individual with bicuspid aortic valve is greater for males versus females.[10]

We recently reported the conditionally dependent relationships between genotypes, gene functions and phenotypes in a large CHD cohort using Bayesian Networks, an Artificial Intelligence (AI) solution.[8] Compared to traditional regression models, Bayesian Networks offer many advantages to risk prediction, as a single network can be used to evaluate multiple outcomes based upon any combination of variables or exposures.[11,12] Moreover, Bayesian Networks provide improved means to quantify the synergistic impacts of demographic, clinical, and genomic variables. In this context, we apply Bayesian Networks in an exploratory, proof-of-principle analysis to explore the landscape of damaging genotypes on neurodevelopmental and linear growth outcomes in the Pediatric Heart Network's (PHN) Single Ventricle Reconstruction (SVR) trial and Infant Single Ventricle (ISV) trial participants.

METHODS

Study population

Participants in the PHN's SVR and ISV trials who had blood or saliva collected for DNA extraction were included in this study. The original study design, descriptive characterization of the cohort, and primary results of the SVR (NCT00115934) and the ISV (NCT 00113087) trials have previously been published.[13-17] Both trials excluded participants with known chromosomal abnormalities or syndromic features at the time of enrollment. Institutional Review Boards for all centers participating in the SVR and ISV trials gave ethical approval for the human subjects research in this work and written

informed consent was obtained. Data Use Agreements between the PHN and the Pediatric Cardiac Genomics Consortium (PCGC) allowed for the de-identified sharing of genetic and limited clinical data between the two consortia. All SVR and ISV participants who had whole exome sequencing (WES) and neurodevelopmental follow-up at 14 months of age were included in this cohort. Sub analyses were performed on those who had complete matching clinical and demographic variables.

Clinical outcomes and Covariates

Outcomes included the Bayley Scales of Infant Development (BSID-II) and length measurement (length for age z-score, LAZ) for both the SVR and ISV trials at 14 months of age. BSID-II subdomain scores included the Psychomotor Developmental Index (PDI) and Mental Developmental Index (MDI), wherein the standardized mean scores are 100. While it is true that MDI and PDI are assessed at the same time, we consider these variables within the Bayesian Networks as both outcomes and risk factors, in order to highlight the ability of our approach to capture and quantify synergistic (non-additive) relationships between variables under study. An advantage of our approach is that the use of a Bayesian network allows employ a single model to explore the impact of highly correlated variables singly or jointly, as outcomes and as risk factors to generate the results shown here. Additional pre- and post-operative clinical and demographic variables were available for a subset of 179 SVR participants, including birthweight, duration of mechanical ventilation following the Stage I palliation surgery, gestational age and US census socioeconomic scores.

In order to optimally discretize continuous variables, we used a grid search to evaluate the performance of the nascent Bayesian Network, using receiver operating characteristic (ROC) curves (Supplementary Figure 1). For each value of MDI, PDI and LAZ, we calculated the joint conditional probability of an individual having low MDI, PDI or LAZ given the presence or absence of other conditions (variables) for each individual in the cohort given its intrinsic values at each condition. These conditional probability distributions were used to evaluate the accuracy of each candidate network to classify an individual as low or high MDI, PDI or LAZ score, allowing us to select the optimal score resulting from the highest area under the ROC (AUC). This approach discovered that a score of 70 for both MDI and PDI resulted in an optimal AUC score (0.97 for PDI and 0.88 for MDI). A score of 70 corresponds to 2 standard deviations below the mean, consistent with clinical intuition for a desirable cut-off. The optimal value for LAZ score using this approach was -1.6 (AUC=0.72). When considering a favorable outcome, we used a threshold of > 100 for MDI and PDI, and > 0 for LAZ. The pre- and post-operative clinical and demographic variables in the smaller subset of 179 SVR participants were less amenable to optimization using this approach. Therefore, we used previously published cutoffs for dichotomization previously associated with outcomes in this population[18,19] as follows: birthweight (< 2500 gram, low), ventilation following the Stage I palliation surgery (> 7 days, prolonged), gestational age (<37 weeks, preterm), and US census socioeconomic score (SES, low score < -0.3, Range -11.5 to 17.9).

Variant annotation

Raw sequencing data in the form of fastq files (Illumina 101 bp paired-end reads) for each individual were aligned to the human genome reference sequence (GRCh37/hg19,

downloaded from http://genome.ucsc.edu) using the Burrows-Wheeler Alignment (BWA) Tool and the MEM algorithm (http://bio-bwa.sourceforge.net/bwa.shtml#13). Variants were called using the best practice protocol of the Genome Analysis Toolkit (GATK) pipeline (http://www.broadinstitute.org/gatk/guide/best-practices). All exomes were jointly-genotyped together with ~190 European individuals (CEU, GBR) from the 1000 Genomes Project. This was followed by Variant Quality Score Recalibration (VQSR), which uses known, high-quality variant sites from HapMap3 and 1000 Genomes to identify and filter potential false positive variant calls. Both variant calling and VQSR were performed separately on INDELs and single nucleotide variants (SNVs). Tranche values were set to 99.5 and 99.0 for SNPs and INDELs, respectively. Variant consequences and annotations were obtained with VEP v.95 utilizing ENSEMBL transcripts version 95 (excluding non-coding transcripts) and selecting the canonical transcript for analysis. Transcript specific prediction for evaluating variant deleteriousness was calculated with VVP,[20] which were also used as input for VAAST.[21] Variants were annotated with ClinVar (20200419) ensuring exact position and base match. Gene-conditions were extracted from the Online Mendelian Inheritance in Man (OMIM) catalogue (2020_07)[22] and the Human Phenotype Ontology (HPO) database (obo file, 2020-08-11).[23] Gene symbols were harmonized using ENSEMBL and HGNC databases controlling for synonymous gene symbols.

AI-based scoring of damaging genetic variants

AI-based prioritization and scoring of candidate disease genes and diagnostic conditions were performed using GEM[24] in the commercially available Fabric Enterprise platform (Fabric Genomics, Oakland, CA). GEM inputs are genetic variant calls in VCF format and proband phenotypes in the form of HPO terms. GEM aggregates inputs from multiple

variant prioritization algorithms (e.g., VAAST and VVP) with genomic and clinical database annotations, using Bayesian means to score and prioritize potentially damaged genes and candidate diseases. The prior probability for the model is based upon known disease associations in the Mendelian conditions databases OMIM and/or HPO with the gene in question. In order to capture a broad category of genes related to heart development, we selected the HPO term Abnormal Heart Morphology (HP:0001627).

GEM's gene scores are Bayes Factors,[25] representing the $\log_{10}$ ratio between the posterior probabilities of two models, summarizing the relative support for the hypothesis that the prioritized genotype damages the gene in which it resides and explains the proband's phenotype versus the hypothesis that the variant neither damages the gene nor explains the proband's phenotype. We used a Bayes factor of $\geq$0.69 to represent a likely deleterious genetic variant, based on a recent retrospective analysis of critically ill newborns whereby a Bayes factor threshold $\geq$0.69 successfully achieved a diagnostic rate of 95% for cases solved using a standard clinical diagnostic approach.[24] Each variant that passed the 0.69 threshold was evaluated using the graph-based adjudication tool Graphite, which constructs a variant graph that permits the exclusion of false positives due to errors in variant calling.[8] Gene penetrance for GEM calculations was set to 0.95 to enforce strict consideration of known dominant and recessive disorders.

Bayesian Network Analyses

Bayesian Network structure was learned using the R package "bnstruct"[26] which provides a Bayesian Information Criterion (BIC)-based exact structure search algorithm.[27] The exact search algorithm explores the entire applicable space of conditional dependencies in order to discover the optimal network structure for the data. Parameter learning for this optimal network and multimorbidity risk calculations were accomplished using the loopy

belief propagation algorithm from the R gRain package.[28] The conditional dependencies in the dataset are presented as a directed acyclic graph. Risk estimates derived from Bayesian Networks are maximum likelihood estimates given the optimal structure under the BIC and uniform prior probability. Relative risk was calculated in the following manner: Probability (A=True | B=True, C=True) / Probability (A=True | B=False, C=False). Confidence intervals (CI, ± 95%) for the mean values for these estimates were obtained by creating 1000 nets from bootstrap replicates of the same data.

Statistical Approach: Differences in testing scores (MDI and PDI) and LAZ for those with and without damaging genetic variants in various gene categories were assessed using independent non-parametric Mann-Whitney U Test, correcting for multiple comparisons using a Bonferroni correction.

RESULTS

Of the original 555 SVR and 230 ISV trial enrollees, survival rates have been previously published.[15,17] Of those that underwent WES, there were 21 individuals who were enrolled in both SVR and ISV trials. For this study, a total of 304 participants underwent WES, neurodevelopmental assessment and length measurement at 14 months of age and form the primary study cohort. Patient-parent trio WES was available for 86 participants. Additionally, the selected pre- and post-operative clinical and demographic variables were available for a subset of 179 SVR participants. Distribution of sex, race, ethnicity, cardiac anatomy, gestational age and birthweight were similar across the various cohorts (Supplementary Table 1). The distribution of neurodevelopmental and growth scores for the cohort undergoing WES (Figure 1) was

similar to that previously reported for the original cohorts and underscore the highly affected nature of these patients.[17,29]

A potentially damaging genetic variant in a gene related to “Abnormal Heart Morphology” was identified in 137 participants (44% of the study cohort), each harboring a mean of 1.8 damaging variants/proband (range 1-6; Supplementary Table 2). Mean MDI, PDI and LAZ scores were not statistically different between participants with and without a damaging genetic variant in this gene list (Figure 2a). Forty-four subjects (14%) harbored one or more damaging genetic variants in a subset of the gene list that is associated with syndromic forms of CHD (as defined by OMIM), including autosomal dominant (N=41), X-linked recessive (N=4), X-linked dominant (N=1) and autosomal recessive (N=1) inheritance modes (Table 1). Eight of these subjects carried variants classified by ClinVar as pathogenic or likely pathogenic. The most commonly damaged genes in the cardiac syndromic subset gene list included *KMT2D* (N=5, Kabuki Syndrome), *NOTCH1* (N=4, Aortic Valve Disease 1), *ANKRD11* (N=3 KBG Syndrome), *MED12* (N=2, Opitz-Kaveggia Syndrome), *SHANK3* (N=2, Phelan-Mcdermid Syndrome), and *SRCAP* (N=2, Floating-Harbor Syndrome). While there was a trend toward lower MDI and PDI scores for participants with *versus* without a damaging variant in the syndromic CHD gene list, the differences were not significant after correction for multiple testing (Figure 2b). We also examined damaging *de novo* genetic variants in the subset of participants with trio WES data (N=86). A damaging *de novo* genetic variant in the HPO-term derived gene list was identified in 10 probands, representing 12% of the cohort with trio WES data (Supplemental Table 3). The small number of *de novo* variants precluded statistical evaluation.

Next, we sought to investigate the conditional probability of particular outcomes, in the context of genetic, demographic and clinical variables, using Bayesian Networks. The Bayesian Network describing the landscape of genetic, neurodevelopmental, growth and demographic variables in our cohort of SV-CHD (Figure 3a) allows for a multitude of queries that support a robust investigation of multiple interactions between conditionally dependent variables. For example, querying the net reveals that a damaging variant in a gene related to abnormal heart morphology increased the risk of a low MDI score by 1.38-fold (1.37 CI-5%, 1.40 CI-95%, N=1000 bootstrap iterations; Figure 3b and Supplementary Table 3). Of the variables under consideration, a low PDI score was the strongest single risk factor for a low MDI score (21.2-fold increase). While female sex was associated with a 1.77-fold increased risk of a low MDI score in isolation, when combined with a low PDI score, the risk of a low MDI was further increased 75.8-fold. These results highlight the ability of Bayesian networks to quantify synergistic effects of variables in combination on an outcome of interest.

The contribution of genetics to PDI score was most impactful when considering predictors of a favorable PDI outcome (Supplementary Table 4). Thus, a participant *without* a damaging variant in a syndromic CHD gene was 4-times more likely to have a normal PDI score (defined as > 100). When combined with normal linear growth (LAZ > 0) and MDI (> 100) scores, the absence of a damaging CHD syndromic genetic variant increased the probability of a normal PDI score by 16.5-fold.

Genetic information modestly impacted poor linear growth, with a damaging variant in a gene related to abnormal heart morphology or a syndromic CHD gene increasing the probability of a low LAZ score by 1.12- and 1.34-fold, respectively (Supplementary Table

4). Additionally, demographic and clinical variables interacted synergistically to increase the probability of a low LAZ score. For example, as isolated variables female sex and low MDI score increased the likelihood of a low LAZ score 1.19-fold and 2.13-fold, respectively. However, in combination, female sex and low MDI further increased the probability of a low LAZ score 6.4-fold.

Next, we sought to determine the impact of pre- and post-operative clinical variables, demographic variables such as SES and sex, and genetics both individually and in combination on neurodevelopmental and linear growth outcomes in a subset of 179 participants for whom complete variables were available. A network describing the landscape of clinical, genetic, demographic, length and neurodevelopmental variable dependencies is depicted in Figure 4. Data elements for this network include premature birth, birthweight, mechanical ventilation duration following the 1st stage palliation surgery and SES. Similar to the larger cohort, we were underpowered to detect the impact of a damaging variant in the syndromic CHD gene list on low MDI, given that only 5 probands harbored a damaging syndromic genetic variant and had an MDI score ≤70. As isolated variables, a damaging genetic variant in a gene related to abnormal heart morphology and prolonged ventilator support following stage I palliative surgery increased the probability of having a low MDI score by 1.92- and 5.76-fold, respectively. However, in combination, these variables acted synergistically to further increase the probability of a low MDI score by 10.5-fold. (Figure 4 and Supplementary Table 5). Likewise, genetic and clinical variables interacted synergistically to predict a favorable neurodevelopmental outcome. For example, the absence of a damaging variant in a known syndromic CHD gene and shorter post-operative ventilator support increased the probability of a normal

MDI score 1.74- and 2.38-fold, respectively, but in combination increased the probability of a good outcome by 59.4-fold.

Similar to the larger cohort, the contribution of genetics to PDI score was most impactful when considering predictors of a favorable PDI outcome (Supplementary Table 5). A participant *without* a damaging variant in a syndromic CHD gene was 1.39-times more likely to have a normal PDI score. However, when combined with a high SES score, the absence of a damaging CHD syndromic genetic variant increased the probability of a normal PDI score by 52.8-fold.

The genetic signal for poor linear growth was smaller in this cohort, in part due to loss of power in the smaller dataset. The strongest single variable contributing to poor linear growth was a low a SES score, increasing the probability of a low LAZ score by 2.13-fold (Figure 4 and Supplementary Table 5). When combined with female sex and a low PDI score, a low SES score increased the probability of a low LAZ score by 7.78-fold.

DISCUSSION

Exploring heterogeneity in outcomes following surgery for complex CHD necessitates models capable of quantifying the complex relationships between clinical variables and genetics, and their combined impacts on divergent outcomes. This is because clinical variables often combine with each other in a conditionally-dependent manner to impact the probability of any given outcome.[9] Here, we apply AI methodologies as a proof-of-principle study of the impact of genetics on neurodevelopmental and linear growth in SV-CHD, using AI-based prioritization of damaging genetic variants and a Bayesian Network approach. Our analyses suggest a modest genetic contribution to

neurodevelopmental outcomes as isolated variables, similar to known clinical predictors. By contrast, we discovered that genetic, demographic, and clinical variables interact synergistically to markedly impact clinical outcomes. These findings highlight the importance of evaluating genetic data in the context of clinical variables, and the need to employ approaches to capture synergistic relationships between them.

Our analyses reveal a modest genetic contribution to poor neurodevelopmental outcomes in probands harboring a damaging variant in abnormal heart morphology genes versus syndromic CHD genes. In the context of predicting a favorable outcome, absence of a damaging variant in the syndromic CHD gene list was more impactful. Although individuals with known syndromic disorders were excluded from enrollment in the SVR and ISV trials, we found that 14% of the SV-CHD cohort harbored a damaging genetic variant in a gene known to cause syndromic CHD, including *KMT2D*, *NOTCH1* and *ANKRD11* as the most common recurring damaged genes. Similar to this cohort, the PCGC also reported a higher-than-expected number of participants with damaging genotypes in syndromic genes, despite a known genetic syndrome as an exclusion criterion.[7] The inherent variable expressivity of cardiac syndromic disorders and/or potential challenges in identifying syndromic features in infancy may, in part, explain these findings. Importantly, our estimates of the impacts of damaging genotypes in syndromic genes (or their absence) are underestimated by the initial exclusion of known syndromic patients. The power of this Bayesian approach is that we can tease apart the contribution of syndromic genes (or other gene classes) to clinical outcomes and thereby predict risk across all patients.

An advantage of Bayesian Networks is the ability to capture and quantify the joint contributions of multiple conditionally dependent variables, all within a single probabilistic graphical model. Traditional techniques, such as regression analyses, have limited ability to capture the conditional dependencies between clinical variables. While mixture and generalized linear models can overcome these weaknesses, they require the design of a new model to address each individual question. Neural nets, while capable of capturing conditional dependencies, lack explainability in that the contributions of different variables on a given output are often unclear. By contrast, Bayesian Networks provide an explainable AI solution.[30]

Our Bayesian Network approach allowed us to start to describe the landscape of clinical, genetic and demographic variables in children with SV-CHD, capturing and quantifying the impacts of synergistic relationships between these variables and outcomes of interest. Genetic variables interacted synergistically with duration of post-operative mechanical ventilation to increase the probability of a poor or favorable MDI score by 10- and 59-fold, respectively. Likewise, genetic variables interacted synergistically with SES to increase the likelihood of a favorable PDI score by 53-fold. Our analyses here were limited to the relatively small number of individuals with complete clinical variables; precluding further exploration of the impact of a damaging genetic variant on pre- and post-operative variables, as well as SES. Expanding the analysis to the wider range of clinical variables associated with this well-phenotyped cohort and increasing the number of genotyped probands will provide greater insight into the genetic contribution to clinical outcomes in SV-CHD. In a broader sense, the current work supplements our efforts using Bayesian Networks as an explainable AI solution for

understanding the joint impacts of diagnoses, medications, and medical procedures on cardiovascular health outcomes.[9]

There are several limitations inherent to this study. For example, DNA collection was not a requirement for participation and thus, sample procurement was inconsistent across the cohort. Moreover, DNA samples were not obtained at enrollment, but rather at subsequent time points in these longitudinally followed cohorts, thereby introducing survival bias. Some DNA samples were collected at a few years of age which would have been well after the 14-month clinical endpoints included in these analyses. Thus, patients who died early may have harbored highly damaging genotypes that are not accounted for in this cohort of infants that survived at least until the 14-month endpoint. The later DNA analysis also highlights the limitations of the original exclusion criteria, as some individuals in SVR were later found (during enrollment in SVR extension studies) were later found to have syndromes, such as Kabuki Syndrome, consistent with our variant analysis. Additionally, we recognize the relatively small sample size of patients limited the ability to include all potential peri-operative, clinical, and socio-economic predictors of outcomes in this proof-of-principal analysis. Future efforts by the PHN consortium to sequence more patients with rich clinical datasets as well by the PCGC to obtain richer clinical data on their sequenced patients, will allow for more robust analyses of the genetic contribution to clinical outcomes in children with complex forms of CHD. The definition of a “damaging” genetic variant is a topic of intense debate.[31] Our rationale for defining a damaging variant was based on a recent study benchmarking the AI tool against clinically solved cases[24]. Stricter criteria for a damaging classification will allow for fewer false positives, but at the expense of false

negatives and *vice versa*. However, the expectation is that false positives will dilute genetic signals, but truly benign variants are *a priori* not positively correlated with negative outcomes. Thus, the magnitudes and statistical significance of the associations between damaging genotypes and clinical variables reported here should be thought of as lower bounds. Finally, the goal of this study was to highlight the advantages of an AI-based Bayesian Networks approach as a proof-of-principle analysis, leveraging existing ISV and SVR clinical and genetic data. In this context, the risk predictions are directly relevant to the study cohort, but cannot be more widely applied without a validation cohort.

Collectively, our findings highlight how genetics, demography and clinical variables interact synergistically to impact neurodevelopmental outcomes and underscore the importance of capturing and quantifying these synergies. We have also demonstrated the utility of Bayesian networks for conditional risk analyses. Moving forward, large-scale collaborations between genetic and clinical consortia will provide more extensive datasets to better understand the synergistic impacts of genetic and clinical variables on clinical outcomes in children with CHD.

DATA AVAILABILITY

The sequencing data used in this analysis may be downloaded, with committee-approved access, from the database of Genotypes and Phenotypes (dbGaP) [www.ncbi.nlm.nih.gov/] (accession numbers phs000571.v5.p2). Clinical data for this cohort is available as a public use dataset (https://www.pediatricheartnetwork.org/login/).

CODE AVAILABILITY

Software used for the analyses may be obtained from the following sites.

GATK: [software.broadinstitute.org/gatk/download/]

VEP: [uswest.ensembl.org/info/docs/tools/vep/script/vep_download.html]

VVP and VAAST: [github.com/Yandell-Lab/VVP-pub]

PHEVOR: [www.yandell-lab.org/software/phevor.html]

GRAPHITE: [github.com/dillonl/graphite]

R: [www.r-project.org] (bnlearn, gRain, bnstruct libraries, pROC package)

GEM: [fabricgenomics.com/fabric-gem]

FUNDING STATEMENT

The clinical data for this project was supported by National Heart, Lung, and Blood Institute (NHLBI) Pediatric Heart Network grants HL068269, HL068270, HL068279, HL068281, HL068285, HL068288, HL068290, HL068292, and HL085057. The genomic data for this project was supported by the NHLBI Pediatric Cardiac Genomics Consortium (UM1-HL098147, UM1-HL128761, UM1-HL098123, UM1-HL128711, UM1-HL098162, U01-HL131003, U01-HL098153, U01-HL098163), the National Center for Research Resources (U01-HL098153), and the National Institutes for Health (R01-GM104390, 1S10OD021644-01A1).

COMPETING INTERESTS

MY is a consultant to Fabric Genomics Inc. and has received consulting fees and stock grants from Fabric Genomics Inc. The remaining authors declare that they have no competing interests relevant to this project.

REFERENCES

1. Menon, S.C. *et al.* Outcome and resource utilization of infants born with hypoplastic left heart syndrome in the Intermountain West. *Am J Cardiol* **110**, 720-7 (2012).
2. Goldberg, C.S. *et al.* Behavior and Quality of Life at 6 Years for Children With Hypoplastic Left Heart Syndrome. *Pediatrics* **144**(2019).
3. Reis, P.M., Punch, M.R., Bove, E.L. & van de Ven, C.J. Outcome of infants with hypoplastic left heart and Turner syndromes. *Obstet Gynecol* **93**, 532-5 (1999).
4. Michielon, G. *et al.* Impact of DEL22q11, trisomy 21, and other genetic syndromes on surgical outcome of conotruncal heart defects. *J Thorac Cardiovasc Surg* **138**, 565-570 e2 (2009).
5. Tabbutt, S. *et al.* Risk factors for hospital morbidity and mortality after the Norwood procedure: A report from the Pediatric Heart Network Single Ventricle Reconstruction trial. *J Thorac Cardiovasc Surg* **144**, 882-95 (2012).
6. Russell, M.W., Chung, W.K., Kaltman, J.R. & Miller, T.A. Advances in the Understanding of the Genetic Determinants of Congenital Heart Disease and Their Impact on Clinical Outcomes. *J Am Heart Assoc* **7**(2018).
7. Jin, S.C. *et al.* Contribution of rare inherited and de novo variants in 2,871 congenital heart disease probands. *Nat Genet* **49**, 1593-1601 (2017).
8. Watkins, W.S. *et al.* De novo and recessive forms of congenital heart disease have distinct genetic and phenotypic landscapes. *Nat Commun* **10**, 4722 (2019).
9. Wesołowski, S. *et al.* An explainable artificial intelligence approach for predicting cardiovascular outcomes using electronic health records. *PLOS Digital Health* **1**(2022).
10. Kong, W.K.F., Bax, J.J., Michelena, H.I. & Delgado, V. Sex differences in bicuspid aortic valve disease. *Prog Cardiovasc Dis* **63**, 452-456 (2020).
11. Pearl, J. *Causality: Models, Reasoning and Inference*, 201-202 (Cambridge University Press, 2000).
12. Pearl, J. Reverend bayes on inference engines: a distributed hierarchical approach. in *Proceedings of the Second AAAI Conference on Artificial Intelligence* 133–136 (AAAI Press, Pittsburgh, Pennsylvania, 1982).
13. Newburger, J.W. *et al.* Transplant-Free Survival and Interventions at 6 Years in the Single Ventricle Reconstruction Trial. *Circulation* (2018).
14. Ohye, R.G. *et al.* Design and rationale of a randomized trial comparing the Blalock-Taussig and right ventricle-pulmonary artery shunts in the Norwood procedure. *J Thorac Cardiovasc Surg* **136**, 968-75 (2008).
15. Ohye, R.G. *et al.* Comparison of shunt types in the Norwood procedure for single-ventricle lesions. *N Engl J Med* **362**, 1980-92 (2010).

16. Hsu, D.T. *et al.* Rationale and design of a trial of angiotensin-converting enzyme inhibition in infants with single ventricle. *Am Heart J* **157**, 37-45 (2009).
17. Hsu, D.T. *et al.* Enalapril in infants with single ventricle: results of a multicenter randomized trial. *Circulation* **122**, 333-40 (2010).
18. Newburger, J.W. *et al.* Early developmental outcome in children with hypoplastic left heart syndrome and related anomalies: the single ventricle reconstruction trial. *Circulation* **125**, 2081-91 (2012).
19. Miller, T. *et al.* The Impact and Interactions of Gestational Age and Weight on Survival and Growth Following the Norwood Procedure: An Analysis from the Pediatric Heart Network Single Ventricle Reconstruction Trial. in *American Heart Association Scientific Sessions* (Anaheim, CA, 2017).
20. Flygare, S. *et al.* The VAAST Variant Prioritizer (VVP): ultrafast, easy to use whole genome variant prioritization tool. *BMC Bioinformatics* **19**, 57 (2018).
21. Hu, H. *et al.* VAAST 2.0: improved variant classification and disease-gene identification using a conservation-controlled amino acid substitution matrix. *Genet Epidemiol* **37**, 622-34 (2013).
22. Online Mendelian Inheritance in Man, OMIM. (McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University, 2022).
23. Köhler, S. *et al.* The Human Phenotype Ontology in 2021. *Nucleic Acids Res* **49**, D1207-D1217 (2021).
24. De La Vega, F.M. *et al.* Artificial intelligence enables comprehensive genome interpretation and nomination of candidate diagnoses for rare genetic diseases. *Genome Medicine* **13**, 153 (2021).
25. Goodman, S.N. Toward evidence-based medical statistics. 1: The P value fallacy. *Ann Intern Med* **130**, 995-1004 (1999).
26. Franzin, A., Sambo, F. & Di Camillo, B. bnstruct: an R package for Bayesian Network structure learning in the presence of missing data. *Bioinformatics* **33**, 1250-1252 (2016).
27. Silander, T. & Myllymaki, P. A simple approach for finding the globally optimal Bayesian network structure. *arXiv preprint arXiv:1206.6875* (2012).
28. Højsgaard, S. Graphical Independence Networks with the gRain Package for R. *Journal of Statistical Software* **46**, 1 - 26 (2012).
29. Newburger, J.W. *et al.* Early developmental outcome in children with hypoplastic left heart syndrome and related anomalies: the single ventricle reconstruction trial. *Circulation* **125**, 2081-91 (2012).
30. Arora, P. *et al.* Bayesian Networks for Risk Prediction Using Real-World Data: A Tool for Precision Medicine. *Value Health* **22**, 439-445 (2019).
31. Eilbeck, K., Quinlan, A. & Yandell, M. Settling the score: variant prioritization and Mendelian disease. *Nat Rev Genet* **18**, 599-612 (2017).

Figure Legends

Figure 1. Outcomes Distribution for the study cohort. Outcome distributions of the Mental Developmental Index (MDI), Psychomotor Developmental Index (PDI) and length for age z-score (LAZ) for the study cohort. Red color denotes values below the optimally derived lower cut-off value (MDI, 70; PDI, 70; LAZ, -1.6) and blue those above the respective cut-off. See Methods for details. N=309 subjects.

Figure 2. Neurodevelopmental and growth outcomes based on presence of a damaging genetic variant: Distribution of Psychomotor Developmental Index (PDI), Mental Developmental Index (MDI), and length for age z-score (LAZ) for those individuals with and without a damaging genetic variant, defined by GEM Bayes Factor $\geq$0.69. (**a**), Box plots of MDI, PDI and LAZ scores in participants with (+) and without (-) a damaging variant in a gene related to the HPO term "Abnormal Heart Morphology"; N=167 participants without damaging genotype, N=137 participants with damaging genotype. None of the comparisons achieved statistical significance after corrections for multiple testing: p=0.94, MDI; p=0.14, PDI; p=0.94, LAZ. (**b**), Box plots of MDI, PDI and LAZ scores in participants with a damaging genotype in a known syndromic CHD gene (N=38) and without (N=167). Center horizontal line, mean value; box, 1 standard deviation; whiskers, 5 and 95th percentiles. None of the comparisons achieved statistical significance after corrections for multiple testing: p=0.68, MDI; p=0.15, PDI; p=0.08, LAZ.

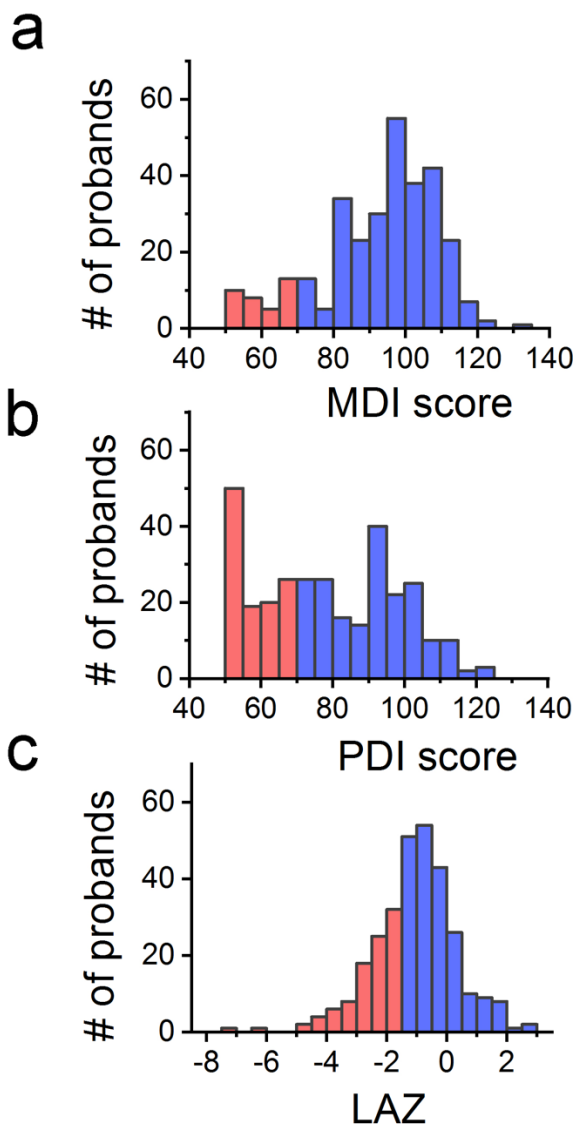
Figure 3. Network describing the conditional dependencies among the genetic, demographic and selected clinical variables in a cohort of children with single ventricle CHD. **a**, Best-fit network for the variables under study. Edges between nodes indicate conditional dependencies. dGV, damaging genetic variant in genes associated with HPO term "Abnormal Heart Morphology". SYNDR, damaging genetic variant in genes associated with syndromic CHD, defined by OMIM. LAZ, length for age Z-score. N=309 participants. **b, c, d,** Forest plots showing relative risk ratios for low MDI ($\leq$ 70), PDI ($\leq$ 70) and LAZ ($\leq$ -1.6) scores in the context of the displayed clinical variables alone or in combination. Solid black line denotes 5 and 95% confidence intervals. If not visible, confidence intervals are within the symbol.

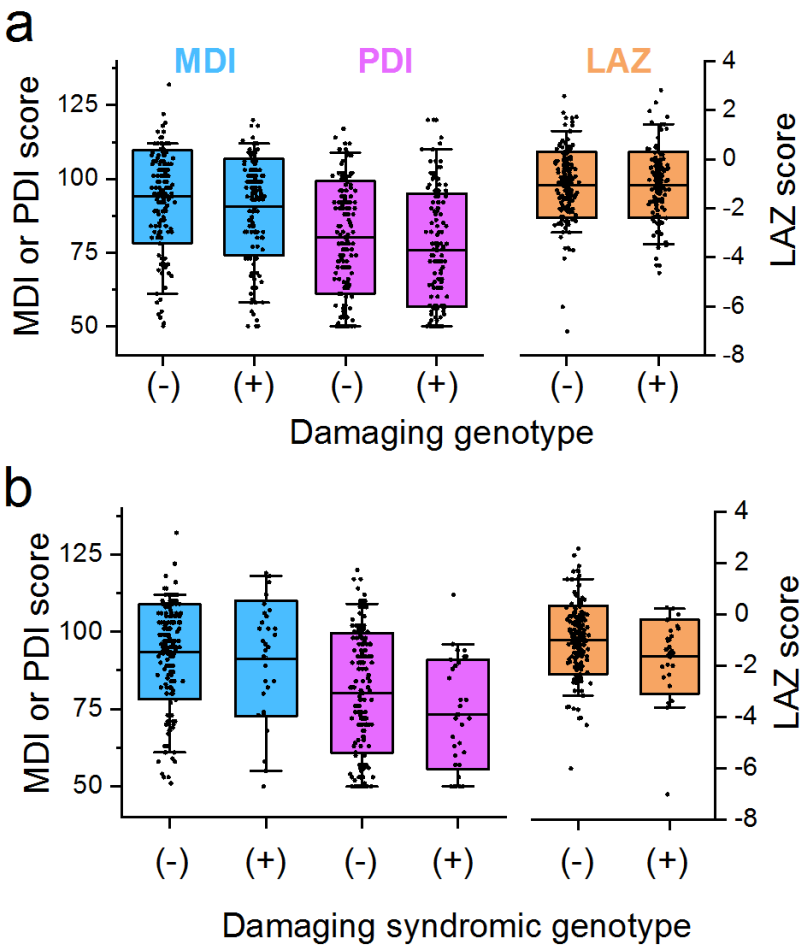
Figure 4. Network describing the conditional dependencies among pre- and post-operative clinical variables, genetics and demographic features in children with single ventricle CHD. a, Best-fit network for the variables under study. Edges between nodes indicate conditional dependencies. BWT, birth weight (<2500 grams, low BWT); VENT, mechanical ventilator duration after 1st stage single ventricle palliation (>7 days, prolonged VENT); preterm, preterm birth (<37 weeks gestation); SES, socioeconomic status score; N=179 participants. **b, c, d,** Forest plots showing relative risk ratios for low MDI ($\leq$ 70), PDI ($\leq$ 70) and LAZ ($\leq$ -1.6) scores in the context of the displayed clinical variables alone or in combination. Solid black line denotes 5 and 95% confidence intervals. If not visible, confidence intervals are within the symbol.

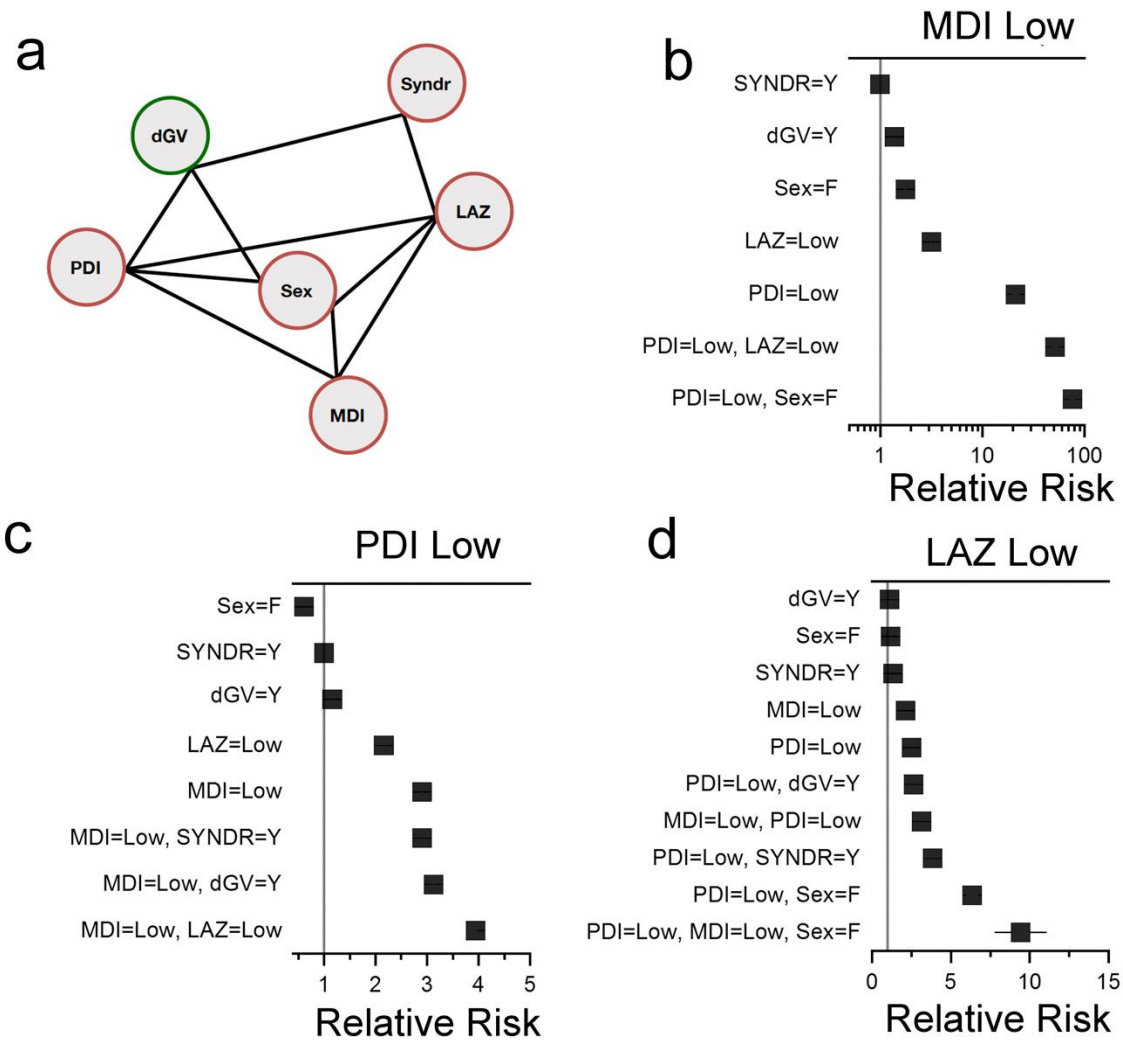
Table 1. Number of probands harboring a damaging variant in a syndromic CHD gene. Damaged genes associated with syndromic CHD (as defined by OMIM) recovered in the study cohort are listed. MIM, phenotype number as listed in OMIM. AD, autosomal dominant; AR, autosomal recessive; XR, X-linked recessive; XD, X-linked dominant.

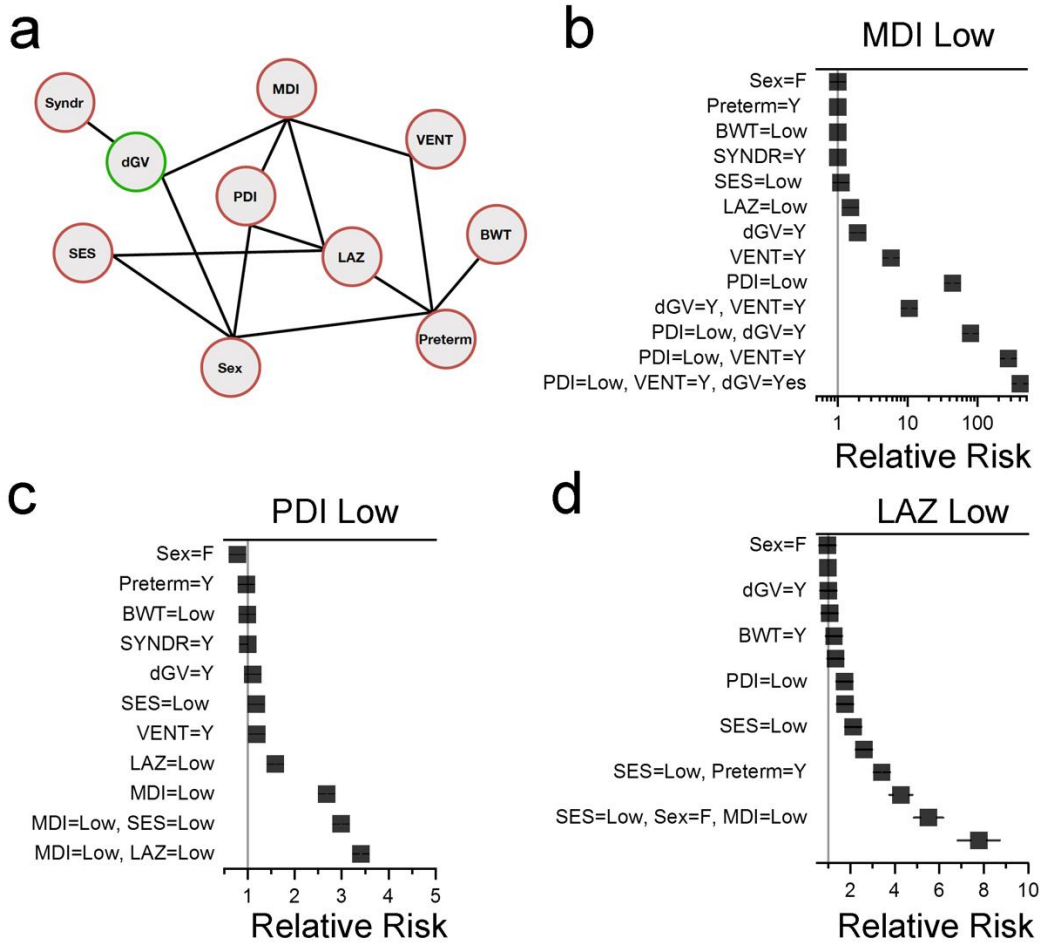
Gene	#	MIM	Disease	Inheritance
KMT2D	5	147920	KABUKI SYNDROME 1	AD
NOTCH1	4	616028	AORTIC VALVE DISEASE 1	AD
ANKRD11	3	148050	KBG SYNDROME	AD
MED12	2	305450	OPITZ-KAVEGGIA SYNDROME	XR
SHANK3	2	606232	PHELAN-MCDERMID SYNDROME	AD
SRCAP	2	136140	FLOATING-HARBOR SYNDROME	AD
AFF4	1	616368	CHOPS SYNDROME	AD
ARHGAP31	1	100300	ADAMS-OLIVER SYNDROME 1	AD
ARID1A	1	614607	COFFIN-SIRIS SYNDROME 2	AD
ATN1	1	618494	CONGENITA HYPOTONIA, EPILEPSY, DEVELOPMENTAL DELAY, AND DIGITAL ANOMALIES	AD
DVL3	1	616894	ROBINOW SYNDROME, AUTOSOMAL DOMINANT 3	AD
CDK13	1	617360	CONGENITAL HEART DEFECTS, DYSMORPHIC FACIAL FEATURES, AND INTELLECTUAL DEVELOPMENTAL DISORDER	AD
CDK8	1	618748	INTELLECTUAL DEVELOPMENTAL DISORDER WITH HYPOTONIA AND BEHAVIORAL ABNORMALITIES	AD
CHD7	1	214800	CHARGE SYNDROME	AD
CREBBP	1	180849	BROAD THUMBS AND GREAT TOES, CHARACTERISTIC FACIES, AND MENTAL RETARDATION	AD
ECE1	1	613870	HIRSCHSPRUNG DISEASE, CARDIAC DEFECTS, AND AUTONOMIC DYSFUNCTION	AD
EED	1	617561	COHEN-GIBSON SYNDROME	AD
EFTUD2		610536	MANDIBULOFACIAL DYSOSTOSIS, GUION-ALMEIDA TYPE	AD
ELN	1	194050	WILLIAMS-BEUREN SYNDROME	AD
EP300	1	613684	RUBENSTEIN-TAYBI TYPE 2	AD
FANCB	1	314390	VACTERL ASSOCIATION, X-LINKED, WITH OR WITHOUT HYDROCEPHALUS	XR
GLI3	1	146510	PALLISTER-HALL SYNDROME	AD
LZTR1	1	616564	NOONAN SYNDROME 10	AD
MEIS2	1	600987	CLEFT PALATE, CARDIAC DEFECTS, AND MENTAL RETARDATION	AD

MYRF	1	618280	CARDIAC-UROGENITAL SYNDROME	AD
NADSYN1	1	618845	VERTEBRAL, CARDIAC, RENAL, AND LIMB DEFECTS SYNDROME 3	AR
NF1	1	193520	WATSON SYNDROME	AD
NSD1	1	117550	SOTOS SYNDROME	AD
OFD1	1	311200	OROFACIODIGITAL SYNDROME I	XD
PBX1	1	617641	CONGENITAL ANOMALIES OF KIDNEY AND URINARY TRACT SYNDROME WITH OR WITHOUT HEARING LOSS, ABNORMAL EARS, OR DEVELOPMENTAL DELAY	AD
POGZ	1	616364	WHITE-SUTTON SYNDROME	AD
SMC1A	1	300590	CORNELIA DE LANGE SYNDROME 2	XR
SOS1	1	610733	NOONAN SYNDROME 4	AD
TBX2	1	618223	VERTEBRAL ANOMALIES AND VARIABLE ENDOCRINE AND T-CELL DYSFUNCTION	AD
TBX5	1	142900	HOLT-ORAM SYNDROME	AD


a
# of probands
0
20
40
60
40
60
80
100
120
140
MDI score
b
# of probands
0
20
40
60
40
60
80
100
120
140
PDI score
c
# of probands
0
20
40
60
-8
-6
-4
-2
0
2
LAZ


a
MDI
PDI
LAZ
MDI or PDI score
LAZ score
125
100
75
50
4
2
0
-2
-4
-6
-8
(-)
(+)
(-)
(+)
(-)
(+)
Damaging genotype
b
MDI or PDI score
LAZ score
125
100
75
50
4
2
0
-2
-4
-6
-8
(-)
(+)
(-)
(+)
(-)
(+)
Damaging syndromic genotype


a
Syndr
dGV
LAZ
PDI
Sex
MDI
b
MDI Low
SYNDR=Y
dGV=Y
Sex=F
LAZ=Low
PDI=Low
PDI=Low, LAZ=Low
PDI=Low, Sex=F
1
10
100
Relative Risk
c
PDI Low
Sex=F
SYNDR=Y
dGV=Y
LAZ=Low
MDI=Low
MDI=Low, SYNDR=Y
MDI=Low, dGV=Y
MDI=Low, LAZ=Low
1
2
3
4
5
Relative Risk
d
LAZ Low
dGV=Y
Sex=F
SYNDR=Y
MDI=Low
PDI=Low
PDI=Low, dGV=Y
MDI=Low, PDI=Low
PDI=Low, SYNDR=Y
PDI=Low, Sex=F
PDI=Low, MDI=Low, Sex=F
0
5
10
15
Relative Risk


a
Syndr
MDI
dGV
VENT
PDI
BWT
SES
LAZ
Preterm
Sex
b
MDI Low
Sex=F
Preterm=Y
BWT=Low
SYNDR=Y
SES=Low
LAZ=Low
dGV=Y
VENT=Y
PDI=Low
dGV=Y, VENT=Y
PDI=Low, dGV=Y
PDI=Low, VENT=Y
PDI=Low, VENT=Y, dGV=Yes
1
10
100
Relative Risk
c
PDI Low
Sex=F
Preterm=Y
BWT=Low
SYNDR=Y
dGV=Y
SES=Low
VENT=Y
LAZ=Low
MDI=Low
MDI=Low, SES=Low
MDI=Low, LAZ=Low
1
2
3
4
5
Relative Risk
d
LAZ Low
Sex=F
dGV=Y
BWT=Y
PDI=Low
SES=Low
SES=Low, Preterm=Y
SES=Low, Sex=F, MDI=Low
2
4
6
8
10
Relative Risk

Supplementary Table 1. Enrollment demographics of each analytic cohort

		SVR and ISV Trial Enrollment Cohorts (n= 785)	Analytic Cohort with MDI/PDI/Length (n=304)	Cohort with full covariates (n=179)	Trios (n=81)
Female Sex (%)		36	35	38	35
Race (%)	White	80	84	88	89
	Black	15	12	9	5
	Asian	2	1	1	2
	Other	3	4	2	4
Ethncity (%)	Hispanic	17	15	14	12
	Non - hispanic	81	84	86	88
	Other	2	1	0	0
HLHS (%)	Yes	81	76	88	85
	No	19	24	12	15
Mean Gestational Age (weeks)		38.0	38.3	38.3	38.2
Mean Birthweight (grams)		3149	3187	3097	3076

Supplementary Table 2. Damaging variants in Abnormal Heart Morphology HPO-derived genes.

ID	GENE	MIM	DISEASE	INHERITANCE	LOCATION	TYPE	CSQ	CLINVAR	GEM Score	SEX	ANCESTRY	AF	GNOMAD
92020048	ABCC6	177850	PSEUDOXANTHOMA ELASTICUM, FORME FRUSTE	ad	16:16248831:G:A	SNV	missense variant	pathogenic_het	2.145	Male	Other	0	gnomad AF=0.00086141
92022650	ABCC6	614473	ARTERIAL CALCIFICATION, GENERALIZED, OF INFANCY, 2	ar	16:16269775:G:A	SNV	missense variant	pathogenic_het	0.956	Male	African	0.0026419	gnomad AF=0.00076492
92022650	ABCC6	614473	ARTERIAL CALCIFICATION, GENERALIZED, OF INFANCY, 2	ar	16:16272764:C:T	SNV	missense variant	pathogenic_het	0.956	Male	African	0.0025241	gnomad_AF=0.00070073
92026925	ABCC6	177850	PSEUDOXANTHOMA ELASTICUM, FORME FRUSTE	ad	16:16256935:G:A	SNV	stop_gained	pathogenic_het	1.907	Female	European(non-Finish)	0.0020086	gnomad_AF=0.0013053
GT04007462	ABCC6	177850	PSEUDOXANTHOMA ELASTICUM, FORME FRUSTE	ad	16:16256935:G:A	SNV	stop_gained	pathogenic_het	2.048	Male	European(non-Finish)	0.0020086	gnomad AF=0.0013053
GT04012012	ABCC6	177850	PSEUDOXANTHOMA ELASTICUM, FORME FRUSTE	ad	16:16255422:C:T	SNV	splice_acceptor_variant	likely_pathogenic_het	2.316	Male	Other	0.0011792	gnomad AF=3.1851e-05
9504000218-0	ACTA2	611788	AORTIC ANEURYSM, FAMILIAL THORACIC 6	ad	10:90701089:A:T	SNV	stop_gained	VUS	1.491	Female	Other	0	no AF in gnomad
GT04013912	ACTG2	155310	VISCERAL MYOPATHY	ad	2:74141940:C:G	SNV	missense_variant	VUS	1.044	Male	European(non-Finish)	0	no AF in gnomad
92123604	ADAR	127400	DYSCHROMATOSIS SYMMETRICA HEREDITARIA	ad	1:154557514:C:A	SNV	missense_variant	VUS	0.754	Female	Finish	0	no AF in gnomad
92021355	AFF4	616368	CHOP5 SYNDROME	ad	5:132220818:A:G	SNV	splice_region_variant,intron_variant	VUS	0.79	Male	European(non-Finish)	0	no AF in gnomad
92020074	ANKRD11	148050	KBG SYNDROME	ad	16:89347773:G:A	SNV	missense_variant	VUS	0.986	Male	European(non-Finish)	0	no AF in gnomad
92026418	ANKRD11	148050	KBG SYNDROME	ad	16:89349242:CTGTT	INDEL	frameshift_variant	pathogenic_het	3.124	Female	European(non-Finish)	0	NA
GT04007303	ANKRD11	148050	KBG SYNDROME	ad	16:89350282:G:A	SNV	missense_variant	VUS	0.731	Male	European(non-Finish)	0	no AF in gnomad
9504000235-0	ANO5	166260	GNATHODIAPHYSEAL DYSPLASIA	ad	11:22291960:TG:T	INDEL	frameshift_variant	pathogenic_het	2.342	Female	European(non-Finish)	0	no AF in gnomad
GT04013282	APOB	144010	HYPERCHOLESTEROLEMIA, FAMILIAL, 2	ad	2:21232182:G:A	SNV	stop_gained	VUS	0.891	Female	Other	0	no AF in gnomad
92023145	ARHGAP31	100300	ADAMS-OLIVER SYNDROME 1	ad	3:119087363:G:A	SNV	splice_region_variant,synonymous_variant	VUS	0.861	Male	Other	0	gnomad AF=3.1865e-05
92121092	ARID1A	614607	COFFIN-SIRIS SYNDROME 2	ad	1:27024018:C:A	SNV	missense_variant	VUS	1.056	Male	European(non-Finish)	0	no AF in gnomad
92024159	ARVCF	NA	NA	ad	22:19965033:T:C	SNV	missense_variant	VUS	0.807	Male	African	0	no AF in gnomad
92027460	ATN1	618494	CONGENITAL HYPOTONIA, EPILEPSY, DEVELOPMENTAL DELAY, AND DIGITAL ANOMALIES	ad	12:7047234:A:C	SNV	splice_region_variant,intron_variant	VUS	0.734	Male	Finish	0	no AF in gnomad
9504000131-0	BCOR	300166	MCOPS2 SYNDROME	xd	X:39937132:G:C	SNV	missense_variant	VUS	0.795	Female	European(non-Finish)	0	gnomad AF=4.5212e-05
GT04007522	BCOR	300166	MCOPS2 SYNDROME	xd	X:39937132:G:C	SNV	missense_variant	VUS	0.776	Female	European(non-Finish)	0	gnomad AF=4.5212e-05
GT04011762	BCOR	300166	MCOPS2 SYNDROME	xd	X:39937132:G:C	SNV	missense_variant	VUS	0.784	Female	European(non-Finish)	0	gnomad AF=4.5212e-05
92120381	BICC1	601331	RENAL DYSPLASIA, CYSTIC, SUSCEPTIBILITY TO	ad	10:60566890:A:G	SNV	missense_variant	VUS	0.711	Male	Other	0	no AF in gnomad
GT04011222	BICC1	601331	RENAL DYSPLASIA, CYSTIC, SUSCEPTIBILITY TO	ad	10:60566890:A:G	SNV	missense_variant	VUS	0.752	Male	Other	0	no AF in gnomad
GT04011733	BTK	300755	AGAMMAGLOBULINEMIA, X-LINKED	xr	X:100612542:A:G	SNV	missense_variant	VUS	0.795	Male	European(non-Finish)	0	no AF in gnomad
92021090	CDH23	617540	PITUITARY ADENOMA 5, MULTIPLE TYPES	ad	10:73565608:A:T	SNV	stop_gained	VUS	0.918	Male	European(non-Finish)	0	no AF in gnomad
GT04011562	CDK13	617360	CONGENITAL HEART DEFECTS, DYSMORPHIC FACIAL FEATURES, AND INTELLECTUAL DEVELOPMENTAL DISORDER	ad	7:40132548:G:T	SNV	stop_gained	VUS	1.129	Female	Other	0	no AF in gnomad
GT04014731	CDK8	618748	INTELLECTUAL DEVELOPMENTAL DISORDER WITH HYPOTONIA AND BEHAVIORAL ABNORMALITIES	ad	13:26927875:A:C	SNV	splice_acceptor_variant	VUS	1.393	Male	European(non-Finish)	0	no AF in gnomad
9504000056-0	CEP290	611134	MECKEL SYNDROME, TYPE 4	ad	12:88477631:G:A	SNV	missense_variant	likely_pathogenic_het	1.255	Male	Other	0	gnomad_AF=6.3739e-05
9504000303-0	CHD7	214800	CHARGE SYNDROME	ad	8:61777576:G:T	SNV	missense_variant,splice_region_variant	VUS	0.839	Male	Other	0	no AF in gnomad
92120189	CHN1	604356	DUANE RETRACTION SYNDROME 2	ad	2:175673698:T:G	SNV	missense_variant	VUS	1.423	Male	European(non-Finish)	0	no AF in gnomad
92024262	CLIP2	NA	NA	ad	7:73790841:C:T	SNV	missense_variant	VUS	0.81	Male	Ashkenazi	0	no AF in gnomad
92026608	CLIP2	NA	NA	ad	7:73774562:C:G	SNV	missense_variant	VUS	0.969	Male	European(non-Finish)	0	no AF in gnomad
GT04006123	CLIP2	NA	NA	ad	7:73731907:G:A	SNV	missense_variant	VUS	0.754	Female	Ashkenazi	0	no AF in gnomad
GT04012081	COG4	618150	SAUL-WILSON SYNDROME	ad	16:70516712:C:A	SNV	stop_gained	pathogenic_het	2.744	Female	Asian	0	gnomad AF=6.3914e-05
GT04010132	COL11A1	228520	FIBROCHONDROGENESIS 1	ad	1:103345432:A:T	SNV	missense_variant	VUS	0.832	Male	European(non-Finish)	0	no AF in gnomad
GT04006123	CR2	240500	IMMUNODEFICIENCY, COMMON VARIABLE, 2	ad	1:207646887:C:T	SNV	splice_region_variant,intron_variant	VUS	0.891	Female	Ashkenazi	0	no AF in gnomad
92029111	CREBBP	180849	BROAD THUMBS AND GREAT TOES, CHARACTERISTIC FACIES, AND MENTAL RETARDATION	ad	16:3778101:G:C	SNV	missense_variant	VUS	0.802	Male	Other	0	no AF in gnomad
GT04011562	CRKL	NA	NA	ad	22:21288342:A:G	SNV	missense_variant	VUS	0.882	Female	Other	0	gnomad_AF=3.1849e-05
92025048	CRYAB	608810	MYOPATHY, MYOFIBRILLAR, 2	ad	11:111781100:T:C	SNV	missense_variant	VUS	0.706	Male	African	0	no AF in gnomad
92120557	CRYAB	608810	MYOPATHY, MYOFIBRILLAR, 2	ad	11:111781100:T:C	SNV	missense_variant	VUS	0.744	Male	African	0	no AF in gnomad
92024159	CTLA4	152700	SYSTEMIC LUPUS ERYTHEMATOSUS	ad	2:204735440:G:A	SNV	missense_variant	VUS	0.883	Male	African	0	no AF in gnomad
92027345	DACT1	617466	TOWNES-BROCKS SYNDROME 2	ad	14:59113613:G:A	SNV	missense_variant	VUS	0.787	Male	European(non-Finish)	0	no AF in gnomad
9504000101-0	DCC	157600	AGENESIS OF CORPUS CALLOSUM	ad	18:50985658:C:T	SNV	missense_variant	VUS	0.732	Female	African	0	no AF in gnomad
GT04012081	DCC	157600	AGENESIS OF CORPUS CALLOSUM	ad	18:50866178:G:A	SNV	missense_variant	pathogenic_het	1.707	Female	Asian	0.0025641	gnomad AF=0.00019116
92021937	DCH51	607829	MITRAL VALVE PROLAPSE 2	ad	11:6654183:G:A	SNV	missense_variant	VUS	0.723	Female	African	0	no AF in gnomad
92021937	DSP	605676	CARDIOMYOPATHY, DILATED, WITH WOOLLY HAIR AND KERATODERMA	ad	6:7565658:G:A	SNV	missense_variant	VUS	0.814	Female	African	0	no AF in gnomad
92123565	DSP	605676	CARDIOMYOPATHY, DILATED, WITH WOOLLY HAIR AND KERATODERMA	ad	6:7584815:GCT:G	INDEL	frameshift_variant	VUS	1.202	Male	European(non-Finish)	0	no AF in gnomad
9504000065-0	DSP	605676	CARDIOMYOPATHY, DILATED, WITH WOOLLY HAIR AND KERATODERMA	ad	6:7569464:G:A	SNV	missense_variant	VUS	0.791	Male	European(non-Finish)	0	no AF in gnomad
9504000325-0	DTNA	604169	LEFT VENTRICULAR NONCOMPACTION 1	ad	18:32462094:G:T	SNV	missense_variant	likely_pathogenic_het	2.65	Female	Other	0	no AF in gnomad
9504000309-0	DVL3	616894	ROBINOW SYNDROME, AUTOSOMAL DOMINANT 3	ad	3:183883971:G:T	SNV	missense_variant	VUS	0.728	Female	European(non-Finish)	0	no AF in gnomad
950400309-0	DVL3	616894	ROBINOW SYNDROME, AUTOSOMAL DOMINANT 3	ad	3:183883971:G:T	SNV	missense_variant	VUS	0.731	Female	European(non-Finish)	0	no AF in gnomad
9504000182-0	ECE1	613870	HIRSCHSPRUNG DISEASE, CARDIAC DEFECTS, AND AUTONOMIC DYSFUNCTION	ad	1:21551931:G:A	SNV	missense_variant,splice_region_variant	VUS	0.738	Male	European(non-Finish)	0	no AF in gnomad
92021393	EED	617561	COHEN-GIBSON SYNDROME	ad	11:85956306:G:C	SNV	missense_variant	VUS	1.448	Male	European(non-Finish)	0	no AF in gnomad
92022282	EFTUD2	610536	MANDIBULOFACIAL DYSOSTOSIS, GUION-ALMEIDA TYPE	ad	17:42931700:C:T	SNV	missense_variant	VUS	1.303	Female	European(non-Finish)	0	no AF in gnomad
9504000092-0	ELN	194050	WILLIAMS-BEUREN SYNDROME	ad	7:73477952:C:T	SNV	splice_region_variant,synonymous_variant	VUS	1.022	Male	European(non-Finish)	0	no AF in gnomad
92021141	EP300	613684	RUBENSTEIN-TAYBI TYPE 2	ad	22:41574281:TGCAA	INDEL	inframe_deletion	likely_pathogenic_het	1.818	Male	Ashkenazi	0	NA
GT04010662	EXT1	133700	EXOSTOSES, MULTIPLE, TYPE I	ad	8:119122874:T:A	SNV	missense_variant	VUS	0.968	Male	European(non-Finish)	0	no AF in gnomad
92025579	EZH2	277590	WEAVER-LIKE SYNDROME	ad	7:148507504:A:G	SNV	splice_region_variant,synonymous_variant	VUS	0.948	Male	African	0	no AF in gnomad
GT04011822	FANCB	314390	VACTERL ASSOCIATION, X-LINKED, WITH OR WITHOUT HYDROCEPHALUS	xr	X:14883175:A:T	SNV	missense_variant	VUS	0.808	Male	Other	0	no AF in gnomad
92120824	FANCG	NA	NA	ad	9:35076529:C:A	SNV	stop_gained	VUS	1.174	Female	Other	0	gnomad AF=3.1863e-05
92020062	FBLN5	614434	CUTIS LAXA, AUTOSOMAL DOMINANT	ad	14:92343929:C:T	SNV	missense_variant	pathogenic_het	1.182	Male	European(non-Finish)	6.48E-05	gnomad AF=3.1851e-05
9504000115-0	FBN1	154700	MARFAN SYNDROME	ad	15:48766740:C:T	SNV	missense_variant	VUS	0.98	Female	European(non-Finish)	0	no AF in gnomad
9504000229-0	FBN2	121050	CONTRACTURAL ARACHNODACTYLY, CONGENITAL	ad	5:127641211:G:T	SNV	missense_variant	VUS	0.872	Female	European(non-Finish)	0	no AF in gnomad
92022422	FGD1	305400	AARSKOG-SCOTT SYNDROME	xr	X:54472639:G:A	SNV	missense_variant	VUS	1.041	Male	European(non-Finish)	0	no AF in gnomad
GT04012142	FHL1	300280	URUGUAY FACIOCARDIOMUSCULOSKELETAL SYNDROME	xr	X:135291407:AG:A	INDEL	frameshift_variant	VUS	1.487	Male	European(non-Finish)	0	no AF in gnomad
9504000218-0	FLRT3	146110	HYPOGONADOTROPIC HYPOGONADISM 7 WITH OR WITHOUT ANOSMIA	ad	20:14306708:A:T	SNV	missense_variant	VUS	1.195	Female	Other	0	no AF in gnomad
92021052	FOS	NA	NA	ad	14:75746823:G:A	SNV	missense_variant	VUS	1.089	Male	European(non-Finish)	0	no AF in gnomad
9504000274-0	GJA1	164200	OCULODENTODIGITAL DYSPLASIA	ad	6:121769125:G:A	SNV	missense_variant	VUS	0.921	Male	Ashkenazi	0	no AF in gnomad
9504000008-0	GLI3	146510	PALLISTER-HALL SYNDROME	ad	7:42005606:G:A	SNV	missense_variant	VUS	0.916	Male	European(non-Finish)	0	no AF in gnomad
9504000367-0	GNE	269921	SIALURIA	ad	9:36218221:G:A	SNV	missense_variant	pathogenic_het	1.412	Male	European(non-Finish)	6.49E-05	gnomad AF=6.3743e-05
92028878	GSN	105120	AMYLOIDOSIS V	ad	9:124086871:T:TGC	INDEL	frameshift_variant	VUS	0.755	Female	Other	0	NA
9504000244-0	GTF2I	194050	WILLIAMS-BEUREN SYNDROME	ad	7:74131206:C:A	SNV	splice_region_variant,intron_variant	VUS	0.699	Male	European(non-Finish)	0	no AF in gnomad
92023258	GTF2IRD1	NA	NA	ad	7:73932531:G:A	SNV	missense_variant	VUS	1.077	Female	European(non-Finish)	0	no AF in gnomad
92029301	GTF2IRD1	NA	NA	ad	7:73922429:C:T	SNV	missense_variant	VUS	0.718	Male	Asian	0	gnomad AF=3.1878e-05
GT04011882	GTF2IRD1	NA	NA	ad	7:73972004:G:A	SNV	missense_variant	VUS	0.85	Male	European(non-Finish)	0	no AF in gnomad
9504000014-0	HCN4	163800	SICK SINUS SYNDROME 2	ad	15:73617719:C:T	SNV	missense_variant	pathogenic_het	2.392	Male	Ashkenazi	0	no AF in gnomad
9504000051-0	HIC1	NA	NA	ad	17:1959641:G:A	SNV	splice_donor_variant	VUS	0.766	Female	African	0.0003445	gnomad_AF=9.5669e-05
92020048	HNRNPA1	615426	AMYOTROPHIC LATERAL SCLEROSIS 20	ad	12:54676358:T:C	SNV	splice_region_variant,synonymous_variant	VUS	0.692	Male	Other	0	no AF in gnomad
92024147	HOXD13	186000	SYNPOLYDACTYLY 1	ad	2:176957822:G:GCG	INDEL	frameshift_variant	VUS	0.801	Male	European(non-Finish)	0	NA
9504000303-0	IFIH1	182250	SINGLETON-MERTEN SYNDROME 1	ad	2:163128888:G:A	SNV	stop_gained	VUS	0.697	Male	Other	0	gnomad AF=6.3723e-05
GT04012532	IKBKG	300636	IMMUNODEFICIENCY 33	xr	X:153780339:T:G	SNV	missense_variant	VUS	1.022	Male	Ashkenazi	0	no AF in gnomad
GT04013282	IL17RD	146110	HYPOGONADOTROPIC HYPOGONADISM 7 WITH OR WITHOUT ANOSMIA	ad	3:57139956:C:T	SNV	missense_variant	likely_pathogenic_het	2.434	Female	Other	0	no AF in gnomad
92021216	INPP5E	213300	CPD4	ar	9:139327608:T:TCC	INDEL	inframe_insertion,splice_region_variant	VUS	0.775	Male	African	0	NA
92021216	INPP5E	213300	CPD4	ar	9:139327605:A:ACC	INDEL	splice_donor_variant	VUS	0.775	Male	African	0	gnomad AF=3.2571e-05
92022282	INS	606176	DIABETES MELLITUS, PERMANENT NEONATAL, WITH NEUROLOGIC FEATURES	ad	11:2181203:C:A	SNV	missense_variant	VUS	0.907	Female	European(non-Finish)	0	no AF in gnomad
9504000194-0	IQSEC2	309530	INTELLECTUAL DEVELOPMENTAL DISORDER, X-LINKED 1	xd	X:53263715:G:A	SNV	missense_variant	VUS	1.356	Female	Other	0	no AF in gnomad
92022371	ITGA8	191830	RENAL HYPODYSPLASIA/APLASIA 1	ad	10:15617485:T:C	SNV	splice_region_variant,intron_variant	VUS	0.987	Male	African	0	no AF in gnomad
92021735	KCNH1	135500	LABAND SYNDROME	ad	1:210857029:T:C	SNV	missense_variant	VUS	1.131	Male	European(non-Finish)	0	no AF in gnomad
9504000024-0	KISS1R	176400	PRECOCIOUS PUBERTY, CENTRAL, 1	ad	19:919929:GCGCGCC	INDEL	inframe_deletion	VUS	1.219	Male	Ashkenazi	0	NA
92027876	KMT2A	605130	WIEDEMANN-STEINER SYNDROME	ad	11:118374319:G:A	SNV	missense_variant	VUS	0.844	Male	European(non-Finish)	0	no AF in gnomad
9504000368-0	KMT2A	605130	WIEDEMANN-STEINER SYNDROME	ad	11:118374568:C:T	SNV	missense_variant	VUS	0.833	Male	European(non-Finish)	0	no AF in gnomad
GT04011592	KMT2A	605130	WIEDEMANN-STEINER SYNDROME	ad	11:118374319:G:A	SNV	missense_variant	VUS	0.892	Male	European(non-Finish)	0	no AF in gnomad
92120189	KMT2D	147920	KABUKI SYNDROME 1	ad	12:49445067:TGA:T	INDEL	frameshift_variant	VUS	1.086	Male	European(non-Finish)	0	no AF in gnomad
9504000274-0	KMT2D	147920	KABUKI SYNDROME 1	ad	12:49420510:ACATT	INDEL	frameshift_variant	likely_pathogenic_het	3.363	Male	Ashkenazi	0	NA
GT04008173	KMT2D	147920	KABUKI SYNDROME 1	ad	12:49446763:A:T	SNV	stop_gained	VUS	1.22	Female	European(non-Finish)	0	no AF in gnomad
GT04010993	KMT2D	147920	KABUKI SYNDROME 1	ad	12:49447337:G:GC	INDEL	frameshift_variant	VUS	1.019	Male	European(non-Finish)	0	no AF in gnomad
GT04012081	KMT2D	147920	KABUKI SYNDROME 1	ad	12:49420288:C:T	SNV	missense_variant	pathogenic_het	2.701	Female	Asian	0	no AF in gnomad
9504000147-0	LAMB3	104530	AMELOGENESIS IMPERFECTA, TYPE IA	ad	1:209799066:G:A	SNV	stop_gained	pathogenic_het	1.941	Male	European(non-Finish)	0.0007127	gnomad_AF=0.00041401
92025175	LDB3	601493	CARDIOMYOPATHY, DILATED, 1C, WITH OR WITHOUT LEFT VENTRICULAR NONCOMPACTION	ad	10:88441257:GCCCA	INDEL	inframe_deletion	VUS	0.698	Male	European(non-Finish)	0	NA
9504000194-0	LETM1	194190	OMIM:194190	ad	4:1824753:G:A	SNV	missense_variant	VUS	1.201	Female	Other	0	no AF in gnomad

92120937	LRP5	259770	OSTEOPOROSIS-PSEUDOGLIOMA SYNDROME	ad	11:68131215:G:A	SNV	splice_region_variant,synonymous_variant	VUS	0.823	Male	European(non-Finish)	0	no AF in gnomad
92023602	LZTR1	616564	NOONAN SYNDROME 10	ad	22:21343948:C:T	SNV	stop_gained	pathogenic_het	2.209	Female	European(non-Finish)	0.0001298	gnomad_AF=9.5669e-05
92026608	MAP1B	618918	PERIVENTRICULAR NODULAR HETEROTOPIA 9	ad	5:71496062:G:A	SNV	missense_variant	VUS	1.086	Male	European(non-Finish)	0	no AF in gnomad
9504000269-0	MAP3K1	613762	46,XY SEX REVERSAL 6	ad	5:56178149:A:T	SNV	missense_variant	VUS	0.722	Male	European(non-Finish)	0	no AF in gnomad
GT04007861	MED12	305450	OPITZ-KAVEGGIA SYNDROME	xr	X:70357766:A:G	SNV	missense_variant	VUS	1.219	Male	European(non-Finish)	0	no AF in gnomad
GT04012732	MED12	305450	OPITZ-KAVEGGIA SYNDROME	xr	X:70344633:C:G	SNV	missense_variant	VUS	0.768	Male	European(non-Finish)	0	no AF in gnomad
9504000194-0	MEIS2	600987	CLEFT PALATE, CARDIAC DEFECTS, AND MENTAL RETARDATION	ad	15:37184453:A:G	SNV	missense_variant	VUS	1.332	Female	Other	0	no AF in gnomad
9504000322-0	MIB1	615092	LEFT VENTRICULAR NONCOMPACTION 7	ad	18:19353605:G:C	SNV	missense_variant	VUS	0.792	Male	European(non-Finish)	0	no AF in gnomad
92120418	MMP21	616749	HETEROTAXY, VISCERAL, 7, AUTOSOMAL	ar	10:127464229:C:A	SNV	missense_variant,splice_region_variant	VUS	0.731	Female	Other	0	no AF in gnomad
92120418	MMP21	616749	HETEROTAXY, VISCERAL, 7, AUTOSOMAL	ar	10:127464224:C:G	SNV	splice_region_variant,intron_variant	VUS	0.731	Female	Other	0	no AF in gnomad
92120367	MPL	254450	MYELOFIBROSIS	ad	1:43804304:C:T	SNV	missense_variant	pathogenic_het	2.962	Male	European(non-Finish)	0	no AF in gnomad
92124125	MSH6	614350	COLORECTAL CANCER, HEREDITARY NONPOLYPOSIS, TYPE 5	ad	2:48030639:A:AC	INDEL	frameshift_variant	pathogenic_het	2.499	Male	European(non-Finish)	0	gnomad_AF=3.2128e-05
9504000200-0	MYBPC3	115197	CARDIOMYOPATHY, FAMILIAL HYPERTROPHIC, 4	ad	11:47364129:C:G	SNV	missense_variant,splice_region_variant	pathogenic_het	1.776	Male	European(non-Finish)	0.0001297	gnomad_AF=6.3747e-05
9504000383-0	MYH11	132900	AORTIC ANEURYSM, FAMILIAL THORACIC 4	ad	16:15810993:A:C	SNV	splice_region_variant,intron_variant	VUS	0.838	Female	European(non-Finish)	0	no AF in gnomad
9504000068-0	MYLK	613780	AORTIC ANEURYSM, FAMILIAL THORACIC 7	ad	3:123454265:T:C	SNV	splice_acceptor_variant	VUS	0.721	Male	European(non-Finish)	0	no AF in gnomad
92020062	MYRF	618280	CARDIAC-UROGENITAL SYNDROME	ad	11:61539309:C:T	SNV	stop_gained	VUS	1.495	Male	European(non-Finish)	0	no AF in gnomad
GT04010762	NADSYN1	618845	VERTEBRAL, CARDIAC, RENAL, AND LIMB DEFECTS SYNDROME 3	ar	11:71202902:G:A	SNV	missense_variant	pathogenic_hom	1.543	Male	European(non-Finish)	0.0023995	gnomad_AF=0.0012108
9504000108-0	NEUROD2	618374	EPILEPTIC ENCEPHALOPATHY, EARLY INFANTILE, 72	ad	17:37762310:C:G	SNV	missense_variant	VUS	1.004	Female	Other	0	no AF in gnomad
9504000200-0	NF1	193520	WATSON SYNDROME	ad	17:29553448:C:T	SNV	splice_region_variant,intron_variant	VUS	0.839	Male	European(non-Finish)	0	no AF in gnomad
9504000072-0	NFKB1	616576	IMMUNODEFICIENCY, COMMON VARIABLE, 12	ad	4:103531776:G:A	SNV	missense_variant	VUS	0.754	Male	European(non-Finish)	0	no AF in gnomad
9504000229-0	NIPBL	122470	BDLS	ad	5:37022498:T:G	SNV	splice_region_variant,intron_variant	VUS	0.774	Female	European(non-Finish)	0	no AF in gnomad
92021521	NONO	300967	MENTAL RETARDATION, X-LINKED, SYNDROMIC 34	xd	X:70514199:G:C	SNV	missense_variant	VUS	0.82	Female	European(non-Finish)	0	no AF in gnomad
9504000084	NOTCH1	616028	AORTIC VALVE DISEASE 1	ad	9:139401755:AC:A	INDEL	splice_donor_variant	VUS	1.049	Male	African	0	no AF in gnomad
9504000111-0	NOTCH1	616028	AORTIC VALVE DISEASE 1	ad	9:139399785:G:T	SNV	stop_gained	VUS	1.13	Female	European(non-Finish)	0	no AF in gnomad
9504000232-0	NOTCH1	616028	AORTIC VALVE DISEASE 1	ad	9:139392011:C:A	SNV	splice_acceptor_variant	VUS	0.983	Male	European(non-Finish)	0	no AF in gnomad
9504000280-0	NOTCH1	616028	AORTIC VALVE DISEASE 1	ad	9:139417442:G:A	SNV	missense_variant	VUS	0.993	Female	European(non-Finish)	0	no AF in gnomad
9504000206-0	NR2F2	615779	CONGENITAL HEART DEFECTS, MULTIPLE TYPES, 4	ad	15:96880841:CA:C	INDEL	frameshift_variant	VUS	1.283	Female	European(non-Finish)	0	no AF in gnomad
9504000235-0	NR5A1	612965	46,XY SEX REVERSAL 3	ad	9:127265367:G:A	SNV	missense_variant	VUS	0.919	Female	European(non-Finish)	0	no AF in gnomad
9504000094-0	NSD1	117550	SOTOS SYNDROME	ad	5:176637132:G:C	SNV	missense_variant	VUS	0.82	Male	Other	0	no AF in gnomad
9504000151-0	NSMF	614838	HYPOGONADOTROPIC HYPOGONADISM 9 WITH OR WITHOUT ANOSMIA	ad	9:140348897:T:TG	INDEL	splice_region_variant,intron_variant	VUS	1.087	Male	European(non-Finish)	0	no AF in gnomad
GT04010722	OFD1	311200	OROFACIODIGITAL SYNDROME I	xd	X:13764945:CA:C	INDEL	frameshift_variant	pathogenic_het	2.12	Male	European(non-Finish)	0	gnomad_AF=0
9504000014-0	OTX2	610125	MICROPHTHALMIA, SYNDROMIC 5	ad	14:57271039:C:T	SNV	missense_variant	VUS	0.948	Male	Ashkenazi	0	no AF in gnomad
GT04008811	PBX1	617641	CONGENITAL ANOMALIES OF KIDNEY AND URINARY TRACT SYNDROME WITH OR WITHOUT HEARING LOSS, ABNORMAL EARS, OR DEVELOPMENTAL DELAY	ad	1:164790855:G:A	SNV	missense_variant	VUS	1.228	Male	Other	0	no AF in gnomad
92021381	PCGF2	618371	TURNPENNY-FRY SYNDROME	ad	17:36895914:C:T	SNV	missense_variant	VUS	0.982	Male	African	0	gnomad_AF=3.1876e-05
92020896	PCSK9	603776	HYPERCHOLESTEROLEMIA, FAMILIAL, 3	ad	1:55517946:G:C	SNV	splice_region_variant,intron_variant	VUS	0.747	Male	Other	0	no AF in gnomad
GT04008261	PIGT	615398	MULTIPLE CONGENITAL ANOMALIES-HYPOTONIA-SEIZURES SYNDROME 3	ad	20:44054311:G:A	SNV	missense_variant	likely_pathogenic_het	0.873	Male	European(non-Finish)	6.48E-05	gnomad_AF=0.00012741
9504000065-0	PKD1	173900	POLYCYSTIC KIDNEY DISEASE 1	ad	16:2147867:A:C	SNV	splice_donor_variant	VUS	0.786	Male	European(non-Finish)	0	no AF in gnomad
92027294	PKP2	609040	ARRHYTHMOGENIC RIGHT VENTRICULAR DYSPLASIA, FAMILIAL, 9	ad	12:32949233:C:T	SNV	splice_acceptor_variant	likely_pathogenic_het	2.997	Male	African	0	no AF in gnomad
92120874	PMS2	614337	COLORECTAL CANCER, HEREDITARY NONPOLYPOSIS, TYPE 4	ad	7:6035203:ACTGT:A	INDEL	frameshift_variant	pathogenic_het	2.518	Male	African	0	gnomad_AF=3.1849e-05
92022282	POGZ	616364	WHITE-SUTTON SYNDROME	ad	1:151377567:C:G	SNV	missense_variant	VUS	0.981	Female	European(non-Finish)	0	no AF in gnomad
92024147	POLR1A	616462	ACROFACIAL DYSOSTOSIS, CINCINNATI TYPE	ad	2:86255740:A:C	SNV	missense_variant	VUS	0.851	Male	European(non-Finish)	0	no AF in gnomad
GT04006052	PPP2R5D	616355	MENTAL RETARDATION, AUTOSOMAL DOMINANT 35	ad	6:42976209:C:T	SNV	missense_variant	VUS	0.744	Male	European(non-Finish)	0	no AF in gnomad
GT04015662	PROK2	610628	HYPOGONADOTROPIC HYPOGONADISM 4 WITH OR WITHOUT ANOSMIA	ad	3:71830676:AT:A	INDEL	frameshift_variant	pathogenic_het	2.253	Male	Other	0	gnomad_AF=6.371e-05
GT04008002	PRRX1	202650	AGNATHIA-OTOCEPHALY COMPLEX	ad	1:170695508:G:A	SNV	missense_variant	VUS	0.768	Male	Other	0	no AF in gnomad
9504000095-0	PTCH2	109400	BASAL CELL NEVUS SYNDROME	ad	1:45295203:C:G	SNV	splice_region_variant,synonymous_variant	VUS	0.729	Female	European(non-Finish)	0	no AF in gnomad
92020961	RET	209880	CENTRAL HYPOVENTILATION SYNDROME, CONGENITAL	ad	10:43607678:C:G	SNV	splice_region_variant,intron_variant	VUS	0.822	Male	European(non-Finish)	0	no AF in gnomad
92029301	RET	209880	CENTRAL HYPOVENTILATION SYNDROME, CONGENITAL	ad	10:43614996:G:A	SNV	missense_variant	pathogenic_het	2.055	Male	Asian	0	gnomad_AF=0.00022333
GT04013812	RET	209880	CENTRAL HYPOVENTILATION SYNDROME, CONGENITAL	ad	10:43619259:G:A	SNV	splice_region_variant,intron_variant	VUS	1.009	Female	European(non-Finish)	0	no AF in gnomad
9504000282-0	RFC2	NA	NA	ad	7:73657477:G:A	SNV	splice_region_variant,synonymous_variant	VUS	1.168	Male	Other	0	gnomad_AF=3.1845e-05
GT04014901	RREB1	NA	NA	ad	6:7248887:G:T	SNV	missense_variant	VUS	1.212	Female	European(non-Finish)	0	no AF in gnomad
GT04011402	RYR1	255310	MYOPATHY, CONGENITAL, WITH FIBER-TYPE DISPROPORTION	ad	19:39013851:C:G	SNV	splice_region_variant,intron_variant	pathogenic_het	2.302	Female	African	0.0004589	gnomad_AF=0.00012741
GT04012532	RYR1	255310	MYOPATHY, CONGENITAL, WITH FIBER-TYPE DISPROPORTION	ad	19:39019237:A:T	SNV	splice_acceptor_variant	VUS	0.777	Male	Ashkenazi	0	no AF in gnomad
GT04013872	SALL4	607323	DUANE ANOMALY WITH RADIAL RAY ABNORMALITIES AND DEAFNESS	ad	20:50401093:A:T	SNV	missense_variant	VUS	0.846	Male	African	0	no AF in gnomad
92120836	SAMHD1	614415	CHILBLAIN LUPUS 2	ad	20:35580045:A:T	SNV	start_lost	VUS	1.371	Female	European(non-Finish)	0	no AF in gnomad
92022042	SCN5A	614022	ATRIAL FIBRILLATION, FAMILIAL, 10	ad	3:38622661:C:A	SNV	missense_variant	pathogenic_het	1.641	Female	European(non-Finish)	6.50E-05	gnomad_AF=3.1922e-05
9504000149-0	SCNN1B	177200	LIDDLE SYNDROME 1	ad	16:23379281:G:A	SNV	splice_donor_variant	VUS	0.733	Male	European(non-Finish)	0	no AF in gnomad
92121129	SEC24C	NA	NA	ad	10:75528652:C:T	SNV	missense_variant	VUS	1.125	Female	European(non-Finish)	0	no AF in gnomad
92120189	SHANK3	606232	PHELAN-MCDERMID SYNDROME	ad	22:51133238:C:T	SNV	missense_variant	VUS	0.734	Male	European(non-Finish)	0	no AF in gnomad
9504000056-0	SHANK3	606232	PHELAN-MCDERMID SYNDROME	ad	22:51160651:C:T	SNV	missense_variant	VUS	0.734	Male	Other	0	no AF in gnomad
9504000182-0	SHH	142945	HOLOPROSENCEPHALY 3	ad	7:155596052:G:A	SNV	missense_variant	VUS	1.14	Male	European(non-Finish)	0	no AF in gnomad
92122373	SIM1	NA	NA	ad	6:100896458:C:A	SNV	missense_variant	VUS	1.447	Female	European(non-Finish)	0	no AF in gnomad
9504000139-0	SIM1	NA	NA	ad	6:100838928:G:A	SNV	missense_variant	VUS	1.184	Female	African	0	no AF in gnomad
GT04008362	SIM1	NA	NA	ad	6:100896458:C:A	SNV	missense_variant	VUS	1.463	Female	European(non-Finish)	0	no AF in gnomad
GT04008021	SLC25A24	612289	GORLIN-CHAUDHRY-MOSS SYNDROME	ad	1:108700128:G:A	SNV	stop_gained	VUS	1.384	Female	European(non-Finish)	0	no AF in gnomad
92020896	SLC25A4	617184	MITOCHONDRIAL DNA DEPLETION SYNDROME 12A (CARDIOMYOPATHIC TYPE), AUTOSOMAL DOMINANT	ad	4:186067983:C:T	SNV	missense_variant	VUS	0.886	Male	Other	0	no AF in gnomad
GT04010132	SLX4	613951	FANCONI ANEMIA, COMPLEMENTATION GROUP P	ar	16:3641280:C:T	SNV	missense_variant	VUS	1.019	Male	European(non-Finish)	0.0016848	gnomad_AF=0.0014013
92025048	SMAD3	613795	LOEYS-DIETZ SYNDROME 3	ad	15:67358499:T:A	SNV	missense_variant	VUS	0.891	Male	African	0	no AF in gnomad
92029147	SMC1A	300590	CORNELIA DE LANGE SYNDROME 2	xr	X:53438775:T:C	SNV	missense_variant	VUS	1.369	Male	European(non-Finish)	0	no AF in gnomad
9504000092-0	SOS1	610733	NOONAN SYNDROME 4	ad	2:39234290:A:T	SNV	missense_variant	VUS	0.898	Male	European(non-Finish)	0	no AF in gnomad
92020947	SPTA1	130600	ELLIPTOCYTOSIS 2	ad	1:158589121:G:A	SNV	missense_variant	likely_pathogenic_het	1.124	Male	European(non-Finish)	0.0030476	gnomad_AF=0.0027721
GT04006072	SPTA1	130600	ELLIPTOCYTOSIS 2	ad	1:158589121:G:A	SNV	missense_variant	likely_pathogenic_het	1.106	Male	European(non-Finish)	0.0030476	gnomad_AF=0.0027721
92020961	SRCAP	136140	FLOATING-HARBOR SYNDROME	ad	16:30750555:G:T	SNV	missense_variant	VUS	0.866	Male	European(non-Finish)	0	no AF in gnomad
92026925	SRCAP	136140	FLOATING-HARBOR SYNDROME	ad	16:30747614:G:T	SNV	missense_variant	VUS	0.88	Female	European(non-Finish)	0	no AF in gnomad
GT04013431	STAT3	147060	HYPER-IgE SYNDROME, AUTOSOMAL DOMINANT	ad	17:40485685:G:A	SNV	splice_region_variant,intron_variant	VUS	0.694	Female	African	0	no AF in gnomad
92022422	SUFU	109400	BASAL CELL NEVUS SYNDROME	ad	10:104352413:A:G	SNV	missense_variant	VUS	0.848	Male	European(non-Finish)	0	no AF in gnomad
GT04013872	SYNE1	612998	EMERY-DREIFUSS MUSCULAR DYSTROPHY 4, AUTOSOMAL DOMINANT	ar	6:152673313:G:A	SNV	missense_variant	VUS	0.813	Male	African	0.0001148	gnomad_AF=3.1892e-05
9504000139-0	TAB2	614980	CONGENITAL HEART DEFECTS, MULTIPLE TYPES, 2	ad	6:149700430:C:A	SNV	missense_variant	VUS	1.32	Female	African	0	no AF in gnomad
9504000182-0	TAB2	614980	CONGENITAL HEART DEFECTS, MULTIPLE TYPES, 2	ad	6:149700172:C:G	SNV	missense_variant	VUS	1.505	Male	European(non-Finish)	0	no AF in gnomad
GT04014901	TAB2	614980	CONGENITAL HEART DEFECTS, MULTIPLE TYPES, 2	ad	6:149700048:C:T	SNV	missense_variant	VUS	0.812	Female	European(non-Finish)	0	no AF in gnomad
9504000026-0	TBL2	NA	NA	ad	7:72984961:G:T	SNV	missense_variant	VUS	0.858	Female	European(non-Finish)	0	NA
9504000218-0	TBX2	618223	VERTEBRAL ANOMALIES AND VARIABLE ENDOCRINE AND T-CELL DYSFUNCTION	ad	17:59485611:T:A	SNV	missense_variant	VUS	1.098	Female	Other	0	no AF in gnomad
92022371	TBX5	142900	HOLT-ORAM SYNDROME	ad	12:114793735:TG:T	INDEL	frameshift_variant	likely_pathogenic_het	3.204	Male	African	0	no AF in gnomad
92123870	TF	209300	ATRANSFERRINEMIA	ar	3:133483071:A:G	SNV	missense_variant	VUS	0.941	Male	European(non-Finish)	0	no AF in gnomad
92123870	TF	209300	ATRANSFERRINEMIA	ar	3:133496032:G:A	SNV	missense_variant	pathogenic_het	0.941	Male	European(non-Finish)	0.0041505	gnomad_AF=0.002357
9504000325-0	TGFBR1	609192	LOEYS-DIETZ AORTIC ANEURYSM SYNDROME	ad	9:101900330:G:A	SNV	missense_variant	VUS	0.746	Female	Other	0	no AF in gnomad
9504000161	TGFBR3	NA	NA	ad	1:92182263:A:G	SNV	splice_region_variant,synonymous_variant	VUS	0.805	Female	Other	0	no AF in gnomad
GT04006421	TK2	117210	SPINOCEREBELLAR ATAXIA 31	ad	16:66551041:G:A	SNV	splice_region_variant,synonymous_variant	VUS	0.915	Male	European(non-Finish)	0	no AF in gnomad
9504000196-0	TNNI3	613690	CARDIOMYOPATHY, FAMILIAL HYPERTROPHIC, 7	ad	19:55668677:C:A	SNV	splice_acceptor_variant	likely_pathogenic_het	3.541	Male	Other	0	no AF in gnomad
9504000065-0	TRIP11	184260	ODONTOCHONDRODYSPLASIA	ad	14:92470821:G:A	SNV	stop_gained	VUS	0.826	Male	European(non-Finish)	0	no AF in gnomad
GT04006322	TRRAP	618454	DEVELOPMENTAL DELAY WITH OR WITHOUT DYSMORPHIC FACIES AND AUTISM	ad	7:98592334:C:G	SNV	missense_variant	VUS	0.719	Male	European(non-Finish)	0	no AF in gnomad
92122664	TSC2	613254	TUBEROUS SCLEROSIS 2	ad	16:2126158:C:G	SNV	missense_variant	VUS	0.764	Female	European(non-Finish)	0	no AF in gnomad
GT04012532	TSC2	613254	TUBEROUS SCLEROSIS 2	ad	16:2129036:G:T	SNV	missense_variant	VUS	0.748	Male	Ashkenazi	0	no AF in gnomad
9504000084	TTR	105210	AMYLOIDOSIS, HEREDITARY, TRANSTHYRETIN-RELATED	ad	18:29178618:G:A	SNV	missense_variant	pathogenic_het	0.934	Male	African	0.017218	gnomad_AF=0.0048432
9504000161	TTR	105210	AMYLOIDOSIS, HEREDITARY, TRANSTHYRETIN-RELATED	ad	18:29172937:G:A	SNV	missense_variant	pathogenic_het	1.996	Female	Other	0	no AF in gnomad
92020505	USP8	219090	PITUITARY ADENOMA 4, ACTH-SECRETING	ad	15:50733549:T:G	SNV	stop_gained	VUS	1.384	Male	European(non-Finish)	0	no AF in gnomad
GT04013752	USP8	219090	PITUITARY ADENOMA 4, ACTH-SECRETING	ad	15:50782004:C:T	SNV	stop_gained	VUS	1.244	Female	Other	0	no AF in gnomad
92121840	WAS	301000	WISKOTT-ALDRICH SYNDROME	xr	X:48547758:C:T	SNV	missense_variant	VUS	1.228	Male	Other	0	no AF in gnomad
9504000282-0	WAS	301000	WISKOTT-ALDRICH SYNDROME	xr	X:48547758:C:T	SNV	missense_variant	VUS	1.32	Male	Other	0	no AF in gnomad
GT04014361	WT1	608978	MEACHAM SYNDROME	ad	11:32417943:C:A	SNV	missense_variant	VUS	1.18	Male	Other	0	no AF in gnomad
92021521	YWHAE	NA	NA	ad	17:1264383:T:C	SNV	splice_region_variant,intron_variant	VUS	0.821	Female	European(non-Finish)	0	no AF in gnomad
GT04008062	YWHAE	NA	NA	ad	17:1265213:C:A	SNV	missense_variant	VUS	1.433	Male	European(non-Finish)	0	no AF in gnomad

Supplementary Table 3. Damaging de novo genetic variants in Abnormal Heart Morphology HPO-derived genes.

PATIENT ID	GENE	LOCATION	TYPE	Consequence	AF	CLINVAR	ANCESTRY	SEX
GT04011562	CRKL	22:21288342:A:G	SNV	missense_variant	0	VUS	Other	Female
GT04010662	EXT1	8:119122874:T:A	SNV	missense_variant	0	VUS	European (non-Finish)	Male
9504000309	FANCM	14:45645935:TCA:T	INDEL	frameshift_variant	0	pathogenic_het	European (non-Finish)	Female
GT04012081	KMT2D	12:49420288:C:T	SNV	missense_variant	0	pathogenic_het	Asian	Female
GT04008173	KMT2D	12:49446763:A:T	SNV	stop_gained	0	VUS	European (non-Finish)	Female
GT04010993	KMT2D	12:49447337:G:GC	INDEL	frameshift_variant	0	VUS	European (non-Finish)	Male
GT04010722	OFD1	X:13764945:CA:C	INDEL	frameshift_variant	0	pathogenic_het	European (non-Finish)	Male
9504000208	OFD1	X:13781856:T:C	SNV	splice_region_variar	0	VUS	European (non-Finish)	Female
9504000014	OTX2	14:57271039:C:T	SNV	missense_variant	0	VUS	Ashkenazi	Male
GT04014901	RREB1	6:7248887:G:T	SNV	missense_variant	0	VUS	European (non-Finish)	Female

AF, allele frequency; VUS, variant of uncertain significance based on Clinvar classification

Supplementary Table 4. Demographic, genetic and clinical contributions to risk of neurodevelopmental and growth outcomes.

Probability Query	Probability	Relative Risk	CI-5%	CI-95%	N
P (MDI=Low)	0.11	-	-	-	36
P (MDI=Low | SYNDR=Y)	0.11	1	1	1	7
P (MDI=Low | dGV=Y)	0.11	1.38	1.37	1.40	20
P (MDI=Low | Sex=F)	0.11	1.77	1.72	1.83	26
P (MDI=Low | LAZ=Low)	0.20	3.17	3.11	3.24	19
P (MD=LowI | PDI=Low)	0.28	21.17	19.19	23.16	32
P (MD=LowI | PDI=Low, LAZ=Low)	0.35	51.28	46.41	56.15	18
P (MDI=Low | PDI=Low, Sex=F)	0.40	75.79	71.04	80.54	23

Probability Query	Probability	Relative Risk	CI-5%	CI-95%	N
P (MDI=High)	0.37				113
P (MDI=High | Sex=M)	0.30	0.82	0.81	0.83	39
P (MDI=High | LAZ=High)	0.46	1.22	1.21	1.23	24
P (MDI=High | dGV=N)	0.37	1.18	1.17	1.18	78
P (MDI=High | SYNDR=N)	0.37	2.17	2.12	2.23	115
P (MDI=High | PDI=High)	0.67	2.41	2.39	2.43	37
P (MDI=High | PDI=High, SYNDR=Y)	0.74	4.03	3.93	4.14	37

Probability Query	Probability	Relative Risk	CI-5%	CI-95%2	N
P (PDI=Low)	0.34	-			115
P (PDI=Low | Sex=F)	0.24	0.61	0.60	0.61	63
P (PDI=Low | SYNDR=Y)	0.35	1	1	1	18
P (PDI=Low | dGV=Y)	0.35	1.16	1.15	1.17	59
P (PDI=Low | LAZ=Low)	0.58	2.16	2.14	2.18	58
P (PDI=Low | MDI=Low)	0.91	2.90	2.88	2.92	32
P (PDI=Low | MDI=Low, SYNDR=Y)	0.91	2.90	2.88	2.92	5
P (PDI=Low | MDI=Low, dGV=Y)	0.92	3.12	3.10	3.15	18
P (PDI=Low | MDI=Low, LAZ=Low)	1.00	3.94	3.91	3.98	18

Probability Query	Probability	Relative Risk	CI-5%	CI-95%	N
P (PDI=High)	0.16	1.00			50
P (PDI=High | Sex=M	0.19	0.97	0.95	0.99	19
P (PDI=High | dGV=N	0.21	1.22	1.21	1.22	30
P (PDI=High | LAZ=High)	0.38	2.09	2.06	2.12	15
P (PDI=High | SYNDR=N)	0.21	3.96	3.61	4.31	49
P (PDI=High | MDI=High)	0.40	5.38	5.26	5.50	39
P (PDI=High | MDI=High, SYNDR=N)	0.42	8.35	7.26	9.43	39
P (PDI =High| MDI=High, LAZ=High)	0.48	10.77	10.40	11.15	9
P (PDI=High | MDI=High, LAZ=High, SYNDR=N)	0.50	16.52	14.26	18.78	9

Probability Query	Probability	Relative Risk	CI-5%	CI-95%	N
P (LAZ=Low)	0.28	-			94
P (LAZ=Low | dGV=Y)	0.32	1.12	1.11	1.12	45
P (LAZ=Low | Sex=F)	0.30	1.19	1.17	1.20	65
P (LAZ=Low | SYNDR=Y)	0.34	1.34	1.32	1.35	18
P (LAZ=Low | MDI=Low)	0.59	2.13	2.10	2.15	19
P (LAZ=Low | PDI=Low)	0.52	2.52	2.49	2.55	58
P (LAZ=Low | PDI=Low, dGV=Y)	0.55	2.64	2.61	2.67	31
P (LAZ=Low | MDI=Low, PDI=Low)	0.64	3.14	3.10	3.18	18
P (LAZ=Low | PDI=Low, SYNDR=Y)	0.76	3.83	3.79	3.88	13
P (LAZ=Low | PDI=Low, Sex=F)	0.53	6.35	5.95	6.75	31
P (LAZ=Low | PDI=Low, MDI=Low, Sex=F)	0.74	9.42	7.80	11.05	15

Probability Query	Probability	Relative Risk	CI-5%	CI-95%	N
P (LAZ=High)	0.18	-			56
P (LAZ=High | Sex=M)	0.15	0.94	0.94	0.95	22
P (LAZ=High | dGV=N)	0.15	1.01	1.01	1.01	32
P (LAZ=High | SYNDR=N)	0.15	1.05	1.04	1.05	49
P (LAZ=High | MDI=High)	0.20	1.36	1.34	1.38	24
P (LAZ=High | PDI=High)	0.29	2.06	2.03	2.09	15
P (LAZ=High | PDI=High, SYNDR=N)	0.28	2.05	2.02	2.09	14

MDI, PDI Low <70; MDI, PDI High >100; LAZ low <-1.6; LAZ High >0

dGV, damaging genetic variant in genes associated with HPO term “Abnormal Heart Morphology”. SYNDR, damaging genetic variant in genes associated with syndromic CHD, defined by OMIM. LAZ, length for age Z-score. CI-5%, Confidence interval 5%; CI-95%, Confidence interval, 95%

Supplementary Table 5. Impacts of pre- and post-operative clinical variables, genetics and demographic features on neurodevelopmental and growth outcomes.

Probability Query	Probability	Relative Risk	CI-5%	CI-95%	N
P(MDI=Low)	0.13	-	-		23
P(MDI=Low | Sex=F)	0.15	1	1	1	17
P(MDI=Low | Preterm=Y)	0.15	1	1	1	2
P(MDI=Low | BWT=Low)	0.15	1	1	1	3
P(MDI=Low | SYNDR=Y)	0.15	1	1	1	5
P(MDI=Low | SES=Low)	0.15	1.11	1.10	1.11	14
P(MDI=Low | LAZ=Low)	0.18	1.51	1.50	1.53	12
P(MDI=Low | dGV=Y)	0.16	1.92	1.86	1.97	14
P(MDI=Low | VENT=Y)	0.23	5.76	5.29	6.24	20
P(MDI=Low | PDI=Low)	0.35	43.54	40.89	46.20	22
P(MDI=Low | dGV=Y, VENT=Y)	0.25	10.49	9.68	11.30	12
P(MDI=Low | PDI=Low, dGV=Y)	0.36	78.62	73.20	84.04	13
P(MDI=Low | PDI=Low, VENT=Y)	0.48	273.04	242.37	303.71	19
P(MDI=Low | PDI=Low, VENT=Y, dGV=Yes)	0.49	401.15	357.88	444.42	11

Probability Query	Probability	Relative Risk	CI-5%	CI-95%
P(MDI=High)				
P(MDI=High | Sex=M)	0.35	1	1	1
P(MDI=High | SES=High)	0.36	1.01	1.00	1.01
P(MDI=High | BWT=High)	0.36	1.07	1.05	1.09
P(MDI=High | Preterm=N)	0.37	1.22	1.16	1.29
P(MDI=High | SYNDR=N)	0.37	1.74	1.49	1.99
P(MDI=High | VENT=N)	0.62	2.38	2.15	2.61
P(MDI=High | PDI=High)	0.77	2.94	2.80	3.07
P(MDI=High | PDI=High, VENT=N)	0.90	5.14	4.77	5.52
P(MDI=High | PDI=High, SYNDR=N)	0.78	5.17	4.33	6.01
P(MDI=High | PDI=High, SYNDR=N, Preterm=N)	0.79	14.96	9.17	20.75
P(MDI=High | SYNDR=N, VENT=N)	0.62	59.42	50.78	68.05
P(MDI=High | PDI=High, SYNDR=N, VENT=N)	0.91	140.27	119.86	160.67

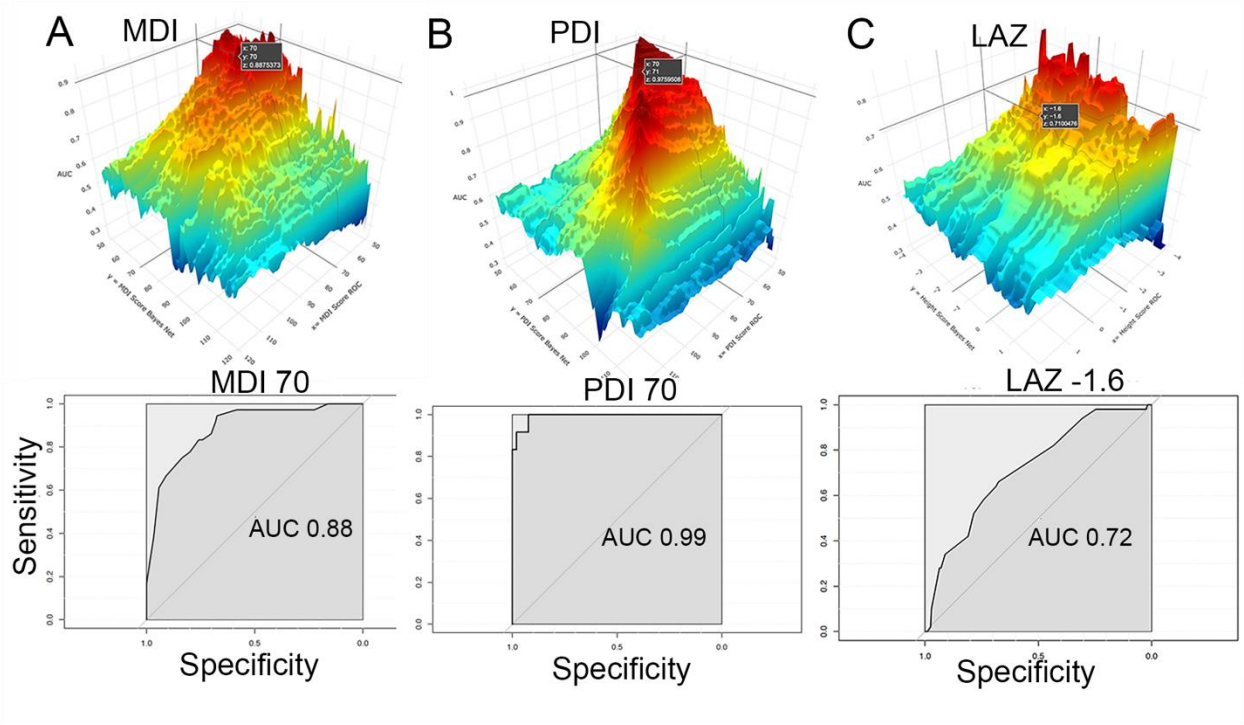
Probability Query	Probability	Relative Risk	CI-5%	CI-95%	N
P(PDI=Low)	0.44				76
P(PDI=Low | Sex=F)	0.36	0.78	0.78	0.79	44
P(PDI=Low | Preterm=Y)	0.44	0.98	0.97	0.98	6
P(PDI=Low | BWT=Low)	0.43	0.99	0.99	0.99	10
P(PDI=Low | SYNDR=Y)	0.43	1.00	1	1.001	9
P(PDI=Low | dGV=Y)	0.45	1.11	1.11	1.12	39
P(PDI=Low | SES=Low)	0.46	1.18	1.18	1.19	44
P(PDI=Low | VENT=Y)	0.49	1.19	1.19	1.20	54
P(PDI=Low | LAZ=Low)	0.51	1.59	1.57	1.61	35
P(PDI=Low | MDI=Low)	0.99	2.68	2.66	2.70	22
P(PDI=Low | MDI=Low, SES=Low)	0.99	2.99	2.96	3.02	14
P(PDI=Low | MDI=Low, LAZ=Low)	0.99	3.41	3.37	3.45	12

Probability Query	Probability	Relative Risk	CI-5%	CI-95%
P(PDI=High)				
P(PDI=High | Sex=M)	0.18	1	1	1
P(PDI=High | LAZ=High)	0.20	1.16	1.13	1.19
P(PDI=High | BWT=High)	0.19	1.25	1.19	1.32
P(PDI=High | Preterm=N)	0.19	1.25	1.19	1.32
P(PDI=High | SYNDR=N)	0.18	1.39	1.26	1.51
P(PDI=High | VENT=N)	0.27	1.75	1.65	1.86
P(PDI=High | SES=High)	0.25	2.31	1.96	2.67
P(PDI=High | MDI=High)	0.39	7.49	5.84	9.14
P(PDI=High | SYNDR=N, SES=High)	0.33	52.76	45.64	59.87
P(PDI=High | MDI=High, SES=High)	0.44	164.93	146.98	182.89

Probability Query	Probability	Relative Risk	CI-5%	CI-95%	N
P(LAZ=Low)	0.35				62
P(LAZ=Low | Sex=F)	0.32	0.97	0.97	0.98	40
P(LAZ=Low | SYNDR=Y)	0.33	1.00	1.00	1.00	9
P(LAZ=Low | dGV=Y)	0.33	1.02	1.02	1.02	28
P(LAZ=Low | VENT=Y)	0.34	1.07	1.07	1.07	42
P(LAZ=Low | BWT=Y)	0.37	1.27	1.26	1.28	12
P(LAZ=Low | MDI=Low)	0.39	1.34	1.33	1.34	12
P(LAZ=Low | PDI=Low)	0.39	1.74	1.72	1.76	35
P(LAZ=Low | Preterm=Y)	0.45	1.76	1.74	1.79	8
P(LAZ=Low | SES=Low)	0.43	2.13	2.10	2.17	42
P(LAZ=Low | SES=Low, BWT=Y)	0.48	2.62	2.57	2.66	7
P(LAZ=Low | SES=Low, Preterm=Y)	0.57	3.42	3.36	3.47	4
P(LAZ=Low | SES=Low, Sex=F)	0.47	4.28	3.76	4.80	24
P(LAZ=Low | SES=Low, Sex=F, MDI=Low)	0.54	5.52	4.86	6.18	8
P(LAZ=Low | SES=Low, Sex=F, PDI=Low)	0.54	7.78	6.83	8.73	13

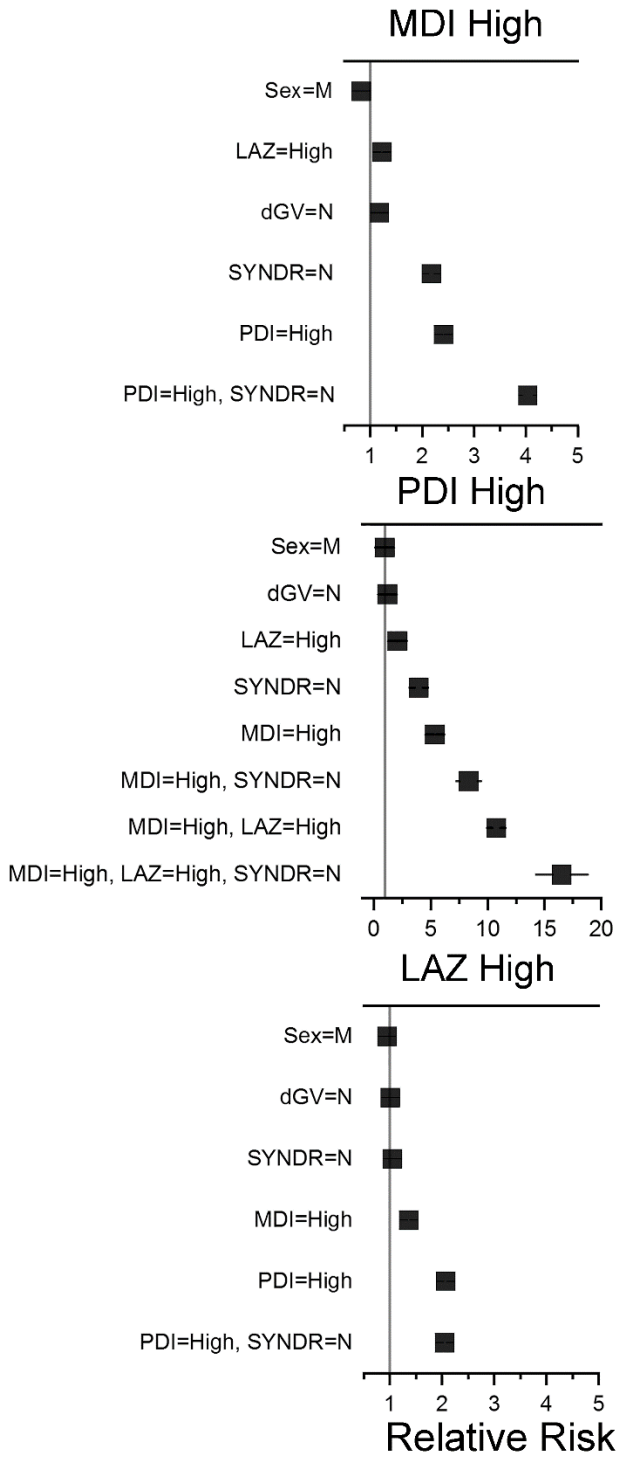
Probability Query	Probability	Relative Risk	CI-5%	CI-95%
P(LAZ=High)				
P(LAZ=High | SYNDR=N)	0.20	1	1	1
P(LAZ=High | MDI=High)	0.20	1	1	1
P(LAZ=High | VENT=N)	0.20	1.01	1.00	1.01
P(LAZ=High | Preterm=N)	0.20	1.09	1.07	1.11
P(LAZ=High | PDI=High)	0.22	1.16	1.13	1.18
P(LAZ=High | BWT=High)	0.20	1.18	1.14	1.22
P(LAZ=High | Sex=M)	0.22	1.30	1.10	1.50
P(LAZ=High | SES=High)	0.27	2.29	2.01	2.57

Supplementary Figure 1



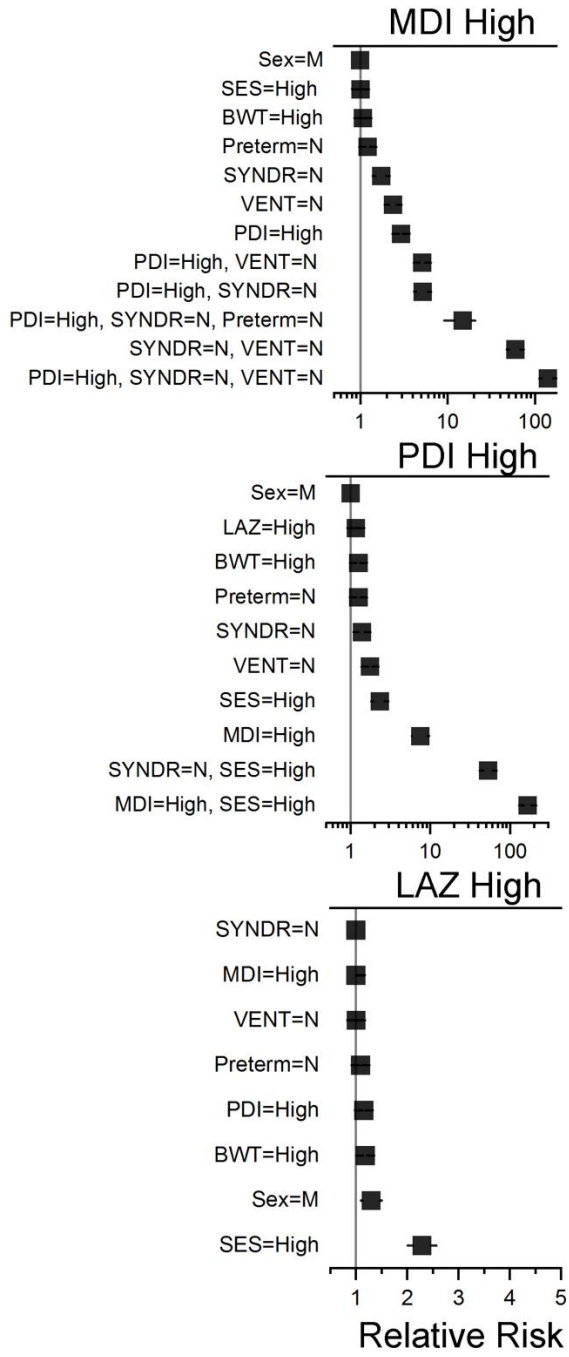
Supplementary Figure 1. Optimal discretization of continuous outcome variables. 3-dimensional heat maps for MDI (**a**), PDI (**b**) and LAZ (**c**) derived from a grid search of the nascent Bayesian Network, maximizing the area under a receiver operating curve (Z-axis), as a function of the posterior probabilities (X-axis) and continuous values (Y-axis) for each outcome variable. The ROC curves for the optimized values of MDI (70), PDI (70) and LAZ (-1.6 are displayed below each heat map, with their respective AUC listed. See text for details.

Supplementary Figure 2



Supplementary Figure 2. Forest plots showing relative risk ratios for high MDI (>100), PDI (>100) and LAZ (>0) scores in the context of the displayed clinical variables alone or in combination. Solid black line denotes 5 and 95% confidence intervals. N=309 participants. If not visible, confidence intervals are within the symbol.

Supplementary Figure 3



Supplementary Figure 3. Forest plots showing relative risk ratios for high MDI (>100), PDI (>100) and LAZ (>0) scores in the context of the displayed clinical variables alone or in combination. 5 and 95% confidence intervals are within the symbols. N=172 participants. If not visible, confidence intervals are within the symbol.