DEVELOPMENT AND OPTIMIZATION OF A CLINICAL SUPPORT ALGORITHM FOR RAPID IDENTIFICATION OF DIAGNOSTIC GERMLINE VARIANTS

Robert J Schuetz[1], Austin A Antoniou[1], David M Gordon[1], Harkness C Kuck[1], Bimal P Chaudhari[1,2] and Peter White[1,2]

[1] Computational Genomics Group, The Institute for Genomic Medicine, Nationwide Children's Hospital, Columbus, Ohio, USA
[2] Department of Pediatrics, College of Medicine, The Ohio State University, Columbus, Ohio, USA

Running title: Probabilistic ranking of genomic variants

Abstract

Congenital disorders are the leading cause of death amongst infants in the U.S. and ultimately effect approximately 8% of the population. Next generation sequencing methods have contributed to increased diagnostic yield in rare disease diagnostics; however, most patients referred to genetics departments still do not receive a diagnosis. By leveraging computational methods, candidate genetic variants can be ranked by likelihood of causing the disease phenotype. LIRICAL, a likelihood ratio algorithm that implements a phenotype and genotype component, outputs probabilities of candidate variants being diagnostic, which is preferable for human interpretation. Natural language processing (NLP) algorithms are capable of identifying phenotype terms in unstructured clinical notes, but the large number of extracted terms overwhelms LIRICAL and compromises accuracy. Here we compare our improved likelihood ratio algorithm, CAVaLRi, and investigate the clinical utility of NLP generated phenotype sets. Novel features of CAVaLRi include limiting inputted phenotype sets to only the most informative terms, incorporating parental genotypes and assigning relative importance by weighting each likelihood ratio component. Genetic sequencing data from an internal cohort (n=611, solved=185) were obtained along with phenotype sets curated by clinical staff. Clinical notes from the electronic health record were passed to ClinPhen, an NLP phenotype extraction algorithm, to generate computational phenotype sets. When passing clinician curated phenotype sets, CAVaLRi significantly outperformed LIRICAL (ROC AUC improved from 0.80 to 0.94, average rank of solved cases improved from 11.4 to 5.7, p=7.97e-16). CAVaLRi accuracy was virtually identical when clinician curated phenotype sets were replaced by ClinPhen generated phenotype sets (ROC AUC remained unchanged at 0.94, average rank of solved cases increase trivially from 5.7 to 5.8, p=0.23). The likelihood ratio paradigm extensions provided by CAVaLRi lead to highly significant gains in diagnostic variant classification accuracy compared to leading variant prioritization algorithms. CAVaLRi stands as the best available computational tool for ensuring diagnostic variants are not overlooked in clinical review.

Introduction

Rare genetic diseases effect approximately 8% of the U.S. population and collectively represent the leading cause of infant mortality[1]. Approximately 80% of rare diseases are thought to be caused by variants in a patient's genome. However, only 25-40% of patients suspected of having a rare disease receive a diagnosis, even after extensive testing[2,3]. Efforts to increase this diagnostic yield have led to the curation of databases, such as Online Mendelian Inheritance in Man (OMIM), that aim to standardize the description of the molecular basis of the thousands of known rare genetic diseases[4]. Gene to disease annotations are available for all OMIM disease entries, allowing clinical genetics teams to infer whether pathogenic variation in a given gene may explain a patient's phenotype.

While advances in exome sequencing (**ES**) and genome sequencing (**GS**) technologies have drastically improved our ability to find disease causing genetic variants, interpreting GS data presents a significant challenge. Efficient prioritization of potentially disease causing variants has remained a challenge given the large number of variants detected in ES analysis[5]. When compared to the human reference genome, an individual's genome sequence may contain anywhere from 3 to 6 million single nucleotide variants (SNVs). Even after filtering variants of low quality and likely benign significance (i.e. single nucleotide polymorphisms (SNPs)), hundreds of possibly disease-related variants remain to undergo manual review. This review process entails variant scientists and clinicians assessing variants for both implied pathogenicity and matching of disease(s) associated with the gene to the patient's clinical features, or phenotypes. This phenotype matching is based on previously published case studies with well characterized pathogenic variants in the gene in which the candidate variant resides. Additionally, variant scientists analyze available parental genome data to ascertain whether candidate diagnostic variants have an inheritance

pattern that is consistent with the disease. For example, a fully penetrant autosomal dominant condition that only requires one disease allele to be present would be discounted if the variant was inherited from an unaffected parent. This process is inherently subjective however, particularly when there is ambiguity as to which elements of a patient's phenotype are genetic in origin. This can result in differing interpretations between clinicians. This emphasizes the need for reproducible, objective approaches to better prioritize diagnostic variants and ensure that high ranking variants are not overlooked in manual review.

Current American College of Medical Genetics (**ACMG**) guidelines consider in silico methods for predicting pathogenicity as secondary evidence at best, and manual review of all available lines of evidence remains the accepted standard[6]. While computational approaches to variant prioritization have been proposed, only a minority allow phenotypic data to be inputted along with genetic variant data[7–9]. These methods have the advantage of modeling the previously described manual phenotype comparison as well as the pathogenicity of the genetic variant. Human phenotypes have been extensively catalogued in the Human Phenotype Ontology (**HPO**), allowing for disease frequency annotation and implicit hierarchical assignment between phenotypes[10]. This mapping from HPO phenotypes to disease probabilities has become increasingly relevant given the independent development of natural language processing (**NLP**) capable of extracting HPO terms from plain text medical records. Distilling signal from these computationally derived phenotype sets would allow for partially automated interpretation of germline variants in settings where HPO phenotype sets are not curated by clinicians or where such manual curation does not scale well.

While leading variant prioritization algorithms report high diagnostic accuracy, the underlying architectures are not without limitations, as certain clinical assessment procedures are

omitted. Here we present The **C**linical **A**ssessment of **Va**riant **L**ikelihood **R**ati**i**os framework, (**CAVaLRi**), that incorporates extensions to the previously reported likelihood ratio paradigm[7] to better identify diagnostic genetic variation. By carefully filtering inputted phenotypes, CAVaLRi is designed to handle larger computational phenotype sets and thus can be combined with NLP algorithms to return accurate results as part of an automated interpretation pipeline. By increasing the accuracy with which diagnostic variants are identified, the burden on clinical teams will be reduced by limiting the number of variants requiring review and ultimately allow for more diagnoses to be returned to patients and their families in less time.

Materials and Methods

The likelihood ratio paradigm for augmented variant prioritization was first implemented in LIRICAL[7], and demonstrated superior performance compared to leading variant prioritization algorithms, namely the Exomiser PhenIX and hiPHIVE models[9]. Likelihood ratios are advantageous, as they are relatively intuitive and can be converted to probabilities given a prior distribution (p) and test results (X), which are highly interpretable for variant scientists and clinicians.

$$\Pr(D \mid \mathrm{X}) = \frac{p * \mathrm{compositeLR}_D}{(1-p) + p * \mathrm{compositeLR}_D} \qquad \text{(Equation 1)}^7$$

Components of the clinical diagnosis workflow can be modeled by separate likelihood ratios that can in turn be multiplied to form a single composite likelihood ratio, as is seen in the LIRICAL definition where a phenotype and genotype likelihood ratio equally contribute to the final probability calculation.

$$\mathrm{compositeLR}_D = \mathrm{phenoLR}_D * \mathrm{genoLR}_D \qquad \text{(Equation 2)}^7$$

CAVaLRi was developed to extend the likelihood ratio framework and address limitations in the LIRICAL approach **(Figure 1)**. The first extension to the framework is the inclusion of filtering

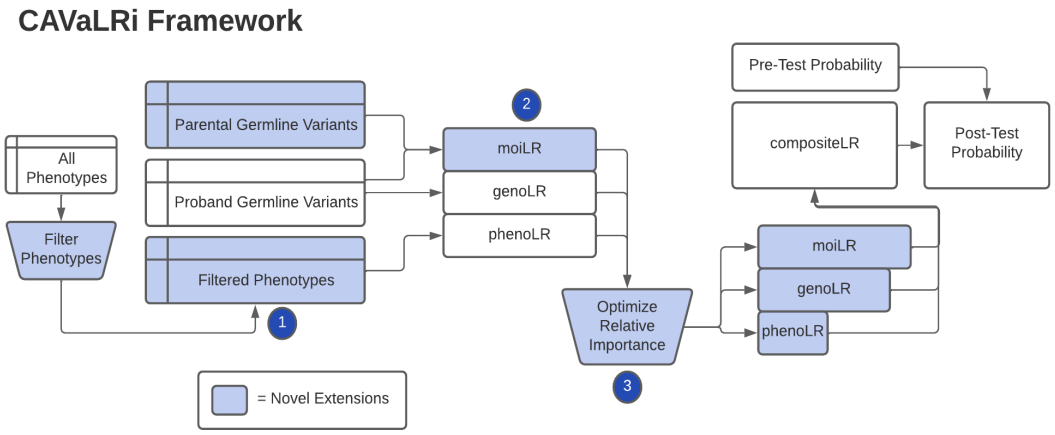

Figure 1 The LIRICAL framework consists of two likelihood ratio components, the genotype likelihood ratio (genoLR) and phenotype likelihood ratio (phenoLR). New components incorporated into the extended framework are indicated in blue, and include filtering phenotype sets (1), adding a mode of inheritance likelihood ratio (moiLR; 2) and rescaling of likelihood ratio components (3).

logic to limit the size of patient phenotype sets to the 15 most informative phenotypes. Phenotype information content (**IC**) was calculated by taking the quotient of the number of genes associated with the query term and the number of genes associated with the parent term(s), prior to applying a negative log function to maps small ratios to large positive values[11]. The resulting IC score is larger when the query term is associated with fewer genes than the parent term(s), indicating a more narrow and informative scope. All given phenotype terms are assigned an IC score, which is then used to sort the phenotypes and remove all but the 15 highest scoring. Once the phenotype set is filtered, phenotype similarity is calculated via a phenotype likelihood ratio[7]. HPO phenotypes are related in a directed acyclic graph structure with "is a" ontology syntax, where descendant terms are more specific, and ascendants are more general (i.e. "large head" is a descendant of "irregularity of head size"). This logic allows for non-exact matching between closely related terms, a necessary feature considering the thousands of HPO entries and limited availability of phenotype disease frequencies. By comparing a patient's phenotypes to known frequencies of HPO

terms for a given disease, similarity can be quantified as a phenotype likelihood ratio (LR_{pheno}). An iterative procedure is carried out to detect this phenotypic disease signal by incrementally enlarging the phenotype set length in calculating the phenotype likelihood ratio. The addition of noise terms to the phenotype set quickly diminishes the similarity signal, with maximum signal peaks occurring in sets no larger than 15 terms in length. Starting with a phenotype set of length five, the is calculated by taking the product of the top five most informative HPO terms (by IC). The next most informative HPO term is then added to the set before recalculating the LR_{pheno} . After incrementing to a set length of 15, the maximum likelihood ratio product is determined from each iterated set length.

$$\mathrm{LR}_{\mathrm{pheno}_D} = \max_{5 \le n \le 15} \prod_{i=1}^{n} LR(\mathrm{HPO}_i)_D \qquad \text{(Equation 3)}$$

Rare diseases largely fall into two mode of inheritance categories, dominant and recessive, indicating the number of pathogenic alleles required to cause disease. Dominant conditions only require one pathogenic allele, while recessive conditions require a pathogenic allele to be present on both the maternal and paternal alleles. Additionally, diseases linked to the X chromosome display differential inheritance patterns between biologically male and female individuals, where females possessing a second functional allele can sometimes suppress a disease phenotype. As previously described, a core component of genetic variant review is to consider the mode of inheritance for the disease in question and determine the genotype of the variant in a patient's parents. Child and parental DNA trio analysis allows for accurate variant phasing and can identify variants that were not inherited from either parent but arose from unfaithful replication of DNA early in embryonic development. These *de novo* variants are highly significant in the rare disease diagnostic setting, but along with other inheritance features, they cannot be detected unless both sets of parental variants are available. Fortunately, parental samples are available in the majority

of pediatric rare disease cases. This affords an opportunity to inform CAVaLRi by algorithmically determining the concordance between the patient's genetic variants, the parents' genetic variants and the annotated mode of inheritance for a given disease. A "mode of inheritance" likelihood ratio (LR_{MOI}) is heuristically defined to return 10^{-1} in the case where variant inheritance does not match the annotated mode of inheritance, 10^{1} if the variant does match and 10^{0} in the case that the status cannot be determined due to lack of parental data availability. In the case that only one parent is present, partial inheritance evidence is modeled by taking the square root of the calculated LR_{MOI} for the parental data that is available. The LR_{MOI} is featured in the composite likelihood ratio that drives CAVaLRi post-test probabilities.

$$\mathrm{LR}_C = \mathrm{LR}_{\mathrm{pheno}} * \mathrm{LR}_{\mathrm{geno}} * \mathrm{LR}_{\mathrm{MOI}} \quad \text{(Equation 4)}$$

The final extension made to the likelihood ratio framework is the addition of exponential constants to differentially scale the likelihood ratios, capturing the relative importance of each modeled clinical component (the LR_{pheno} constant is fixed to 1).

$$\mathrm{LR}_C = \mathrm{LR}_{\mathrm{pheno}} * \mathrm{LR}_{\mathrm{geno}}^{c_1} * \mathrm{LR}_{\mathrm{MOI}}^{c_2} \quad \text{(Equation 5)}$$

The parameterized formula was then subjected to an optimization procedure to find constant values (c_1, c_2) that minimize a pair of loss functions. The first loss function aims to maximize the classification accuracy of candidate variants as diagnostic or non-diagnostic. This is achieved by calculating the C-statistic, or the area under the receiver operating curve, where larger values indicate more accurate binary classification. The range of this function spans from 0.5, the value when the classifier is performing as well as random chance, to 1.0, the value in the case that perfect classification occurs. The C-statistic range was mapped to a [0,1] range for comparability with the second loss function and inverted to change the optimization from maximization to minimization. The second loss function is a normalized Wilcoxon signed rank test statistic that ranges from [0,1]

and captures how much lower diagnostic variants in true positive cases are ranked amongst other non-diagnostic candidate variants. Larger Wilcoxon statistics correlated with a larger difference in rank compared to some baseline ranking method, LIRICAL in our case. This function was also inverted to change the optimization to a minimization problem. Based on the use case, one of these accuracy metrics may be more important in model training. To account for this, we introduce an α variable that can be set to weight one loss function more or less than the other.

$$L = \alpha * L_C + (1 - \alpha) * L_W \quad \text{(Equation 6)}$$

Models were generated for $\alpha \epsilon \{0, 1, 0.5\}$ to provide pretrained models whose scalars optimize binary classification accuracy, average rank of diagnostic variant in solved cases and an equally weighted combination of the two, respectively. To combat against overfitting, fivefold cross validation was employed to find average optimized scalar values. A greedy search optimization approach was employed starting from the origin intersection of the loss function and stepping in the direction of steepest gradient descent. The constant values were constrained to not exceed values larger than 10, and step length was set to 1, constricting the constant values to integer space.

Results

After implementing these extensions, electronic health records and ES data from 611 patients suspected of rare disease (185 diagnosed) were curated from patients evaluated at Nationwide Children's Hospital's Institute for Genomic Medicine (IGM). Manually curated phenotype sets were obtained by querying a locally developed web-based user interface designed to assign phenotype terms to cases. Computationally derived phenotype sets were generated by passing all pre-diagnostic medical records as plain text to ClinPhen, an NLP algorithm leveraging synonym lookup to extract HPO terms[12]. Genetic variants and manually curated phenotype sets for each case were passed to LIRICAL and incrementally extended versions CAVaLRi to assess performance.

After optimization, integer constant values of 4 and 5 were found to be optimal in the equally weighted loss function and were assigned as exponential scalars to the LR_{geno} and LR_{MOI}, respectively. Gains in accuracy were accumulated with the addition of each extension, ultimately resulting in a C-statistic increase from 0.80 to 0.94 (**Figure 2A**). Average rank of diagnostic variants in solved cases decreased from 11.4 to 5.7 (p = 7.97e-16), and the percentage of cases where the diagnostic variant was placed in first position increased from 21% to 58% (**Figure 2B**).

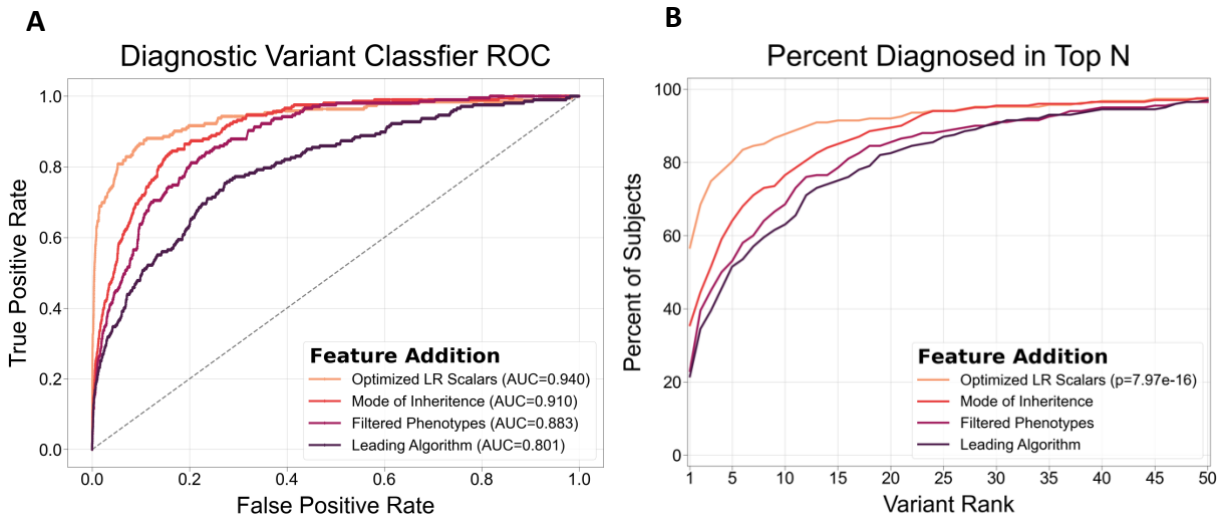

Figure 2 Cumulative gains in accuracy observed with each framework extension LIRICAL, proven to be the leading variant prioritization algorithm, was chosen as the baseline comparator. Extension features were added one at a time and displayed incremental gains in algorithmic accuracy. The accuracy metrics consist of the C-statistic, or area under the ROC curve (**A**) and p-value associated with the one-sided Wilcoxon test statistic (**B**). An equally weighted loss function between these accuracy metrics was used to find optimal likelihood ratio scalar values.

As previously described, noisy HPO terms included in phenotype sets quicky diminish LIRICAL's capacity to prioritize likely diagnostic variants. The phenotype filtering extension of CAVaLRi hedges against phenotypic noise by only considering the most informative phenotypes and by considering the subset of phenotypic terms which make the strongest argument in favor of each diagnosis. Computationally derived phenotype sets include significantly more HPO terms compared to manually curated phenotype sets (**Figure 3A**). These large phenotype sets drive LIRICAL posttest probabilities to infinitesimally small values due to the penalization of terms that are not related to disease but are incidentally noted in a patient chart, rendering the computationally

derived phenotype sets uninformative. To test whether CAVaLRi can detect signal in these large sets, outputted post-test probabilities were calculated when passing computationally generated HPO sets versus clinician curated HPO sets, which have already demonstrated diagnostic utility. Regardless of the phenotype set source, CAVaLRi output was equally accurate according to both the C-statistic and Wilcoxon signed rank test (**Figure 3B**). This result suggests that the phenotype filtering extension to the likelihood interpretation framework is effectively limiting the noise that would otherwise be distorting the diagnostic likelihood signal.

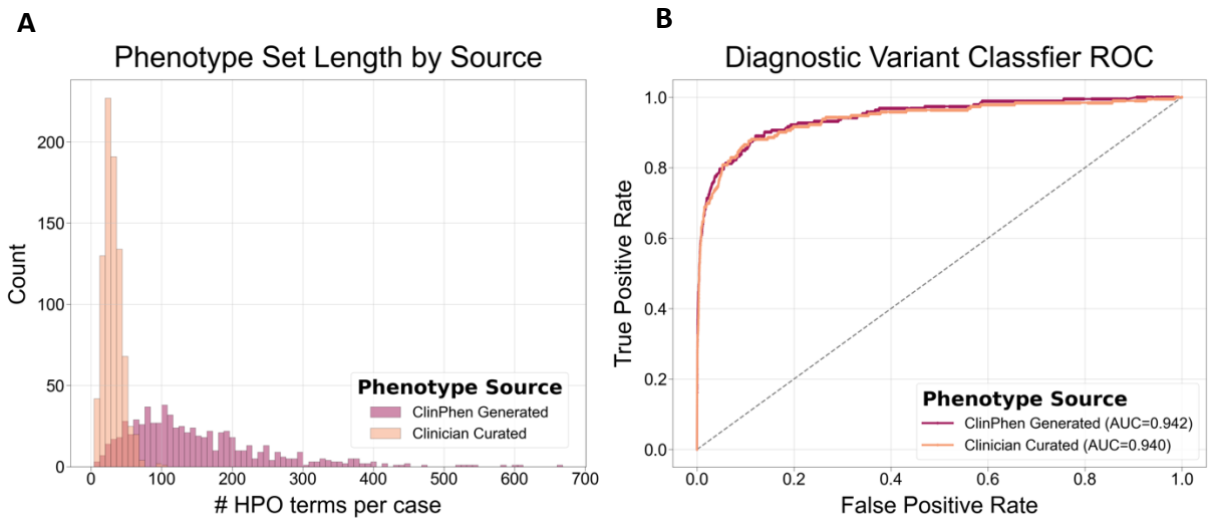

Figure 3 Computational phenotypes prove equally informative in identifying and ranking diagnostic variants LIRICAL struggles to return meaningful results when large phenotype sets are inputted. This is of concern, given that computationally generated phenotype sets from ClinPhen and other NLP algorithms are much larger compared to clinician curated phenotype sets (**A**). Binary classification accuracy, as measured by area under the receiver operator characteristic is not significantly different when computationally derived phenotypes are passed to CAVaLRi versus manually curated phenotype sets (**B**).

Discussion

The reimagined likelihood ratio framework described in CAVaLRi was more effective in classifying genetic variants as diagnostic and was highly effective at placing known diagnostic variants at the top of a sorted candidate gene lists. When compared to LIRICAL, CAVaLRi is the most reliable computational tool available to protect against diagnostic variants being omitted from clinical reports. In cases where the diagnostic variant is identified, turnaround times can expect to

be reduced considering that the average number of variants requiring review has been halved from eleven to five. Regarding sources of improved performance, incremental gains in accuracy were seen in both accuracy metrics as each extension was integrated. The relatively large LR_{MOI} scalar and marked improvement in accuracy upon initial addition to the likelihood framework suggests that the mode of inheritance component is indeed informative and should be integrated when parental variant data is available. The largest gain in accuracy was observed after the optimization procedure, reaffirming that each likelihood ratio has differential importance in determining the composite likelihood ratio. The relatively high scalar values assigned to the LR_{geno} and LR_{MOI} suggest a paradigm where these features are more important than the LR_{pheno}. This finding would suggest that less credence should be placed in phenotypic features versus genotypic features in determining a genetic diagnosis in rare disease. The concept of relative importance of clinical diagnostic workflow components should also be considered in the development of future variant prioritization algorithms.

Filtering phenotypes not only significantly improved accuracy, but also allowed for the extraction of phenotype signal from otherwise noisy sets, specifically computationally derived phenotype sets. This result provides a proof of concept that phenotype curation, a time-consuming and subjective task, may be a candidate for automation. Identical CAVaLRi output accuracy between the two phenotype sources confirms that the reproducible computational methods can be leveraged in a fraction of the time without sacrificing accuracy when compared to manually reading the clinical record and inputting the relevant phenotypes into a user interface. This is also important when considering that not all treatment centers have the resources available to devote to manual phenotype extraction from the medical records. ClinPhen is an open-source algorithm that automatically converts clinical notes into a prioritized list of patient phenotypes with high-

accuracy and speed. Given its ease of implementation, ClinPhen can be combined with CAVaLRi to serve as an automated variant prioritization pipeline.

At the time of writing, no germline variant prioritization algorithm has been proposed that is trained on data including undiagnosed cases suspected of rare disease. This gap in the literature is significant, as enriched training sets of true positive cases are not representative of cases in practice. Our training cohort has not been enriched other than to include subjects who are suspected of rare disease and receive ES analysis, leading to results that are likely to be more generalizable. C-statistics should continue to be used in measuring the accuracy of computational approaches to ensure that true negative variants are identified, this practice is also novel in the setting of germline variant prioritization. CAVaLRi post-test probabilities may prove useful as input to machine learning models aiming to classify patients as diagnosable or non-diagnosable. The CAVaLRi phenotype likelihood ratio specifically could potentially be used prior to ordering ES to identify patients who may have a genetic diagnosis .

Given the low diagnostic yield observed in rare diseases, it can be assumed that there remains an undiagnosed majority in cases of suspected rare disease. Reinterpretation efforts have been extensively documented as highly successful, returning a diagnosis for 5-20% of reanalyzed patients depending on the amount of time between initial consult and ordering of reanalysis[13–17]. When reanalyzing a historical undiagnosed cohort, each case will require manual review of hundreds of variants, an effort that most centers do not have the resources to achieve. CAVaLRi offers a method to sort and prioritize variants, ensuring that clinicians arrive at meaningful findings without the need to manually review every variant that passes filtering. With this functionality in mind, CAVaLRi has been designed to accept cohorts in cases of multiple subjects and ultimately return summary files indicating the most likely diagnostic variants across all individuals.

While these results are promising, certain limitations must be addressed. The only training data available at the time of analysis was internal NCH IGM patient data. To ensure generalizability, the model should ideally be validated on an external cohort that includes subjects who do not harbor a genetic diagnosis. Regarding the scaling of likelihood ratios, the post-test probability deviates from the traditional post-test probability formula. To combat the skewing of post-test probabilities and potential loss of interpretability, the positive predictive value can be used in place of the post-test probability. Heuristics set forth by LIRICAL in calculating phenotype and genotype likelihood ratios have deficiencies that future iterations of CAVaLRi aim to address, including the source of background phenotype frequencies and increasing the resolution with which genetic variant frequencies are defined. Finally, other variant prioritization algorithms have been released, however as previously stated, none of the algorithms have been tested on a cohort containing true negative cases. Before definitively stating that LIRICAL is the leading algorithm, and as follows the best baseline comparator, the negative predictive value of other candidate algorithms should be determined.

CAVaLRi will be published under an open-source license to ensure that any interested institution can perform optimization with their own cohorts and implement their own version of the algorithm in their respective diagnostic workflows.

References

1. Almli LM, Ely DM, Ailes EC, et al. Infant Mortality Attributable to Birth Defects — United States, 2003–2017. *MMWR Morb Mortal Wkly Rep*. 2020;69(2):25-29. doi:10.15585/mmwr.mm6902a1

2. Wright CF, FitzPatrick DR, Firth HV. Paediatric genomics: diagnosing rare disease in children. *Nat Rev Genet*. 2018;19(5):253-268. doi:10.1038/nrg.2017.116

3. Bavisetty S, Grody WW, Yazdani S. Emergence of pediatric rare diseases: Review of present policies and opportunities for improvement. *Rare Dis*. 2013;1(1):e23579. doi:10.4161/rdis.23579

4. Amberger JS, Bocchini CA, Schiettecatte F, Scott AF, Hamosh A. OMIM.org: Online Mendelian Inheritance in Man (OMIM®), an online catalog of human genes and genetic disorders. *Nucleic Acids Res*. 2015;43(D1):D789-D798. doi:10.1093/nar/gku1205

5. on behalf of the Medical Genome Initiative, Marshall CR, Bick D, et al. The Medical Genome Initiative: moving whole-genome sequencing for rare disease diagnosis to the clinic. *Genome Med*. 2020;12(1):48. doi:10.1186/s13073-020-00748-z

6. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17(5):405-424. doi:10.1038/gim.2015.30

7. Robinson PN, Ravanmehr V, Jacobsen JOB, et al. Interpretable Clinical Genomics with a Likelihood Ratio Paradigm. *Am J Hum Genet*. 2020;107(3):403-417. doi:10.1016/j.ajhg.2020.06.021

8. Li Q, Zhao K, Bustamante CD, Ma X, Wong WH. Xrare: a machine learning method jointly modeling phenotypes and genetic evidence for rare disease diagnosis. *Genet Med*. 2019;21(9):2126-2134. doi:10.1038/s41436-019-0439-8

9. Smedley D, Jacobsen JOB, Jäger M, et al. Next-generation diagnostics and disease-gene discovery with the Exomiser. *Nat Protoc*. 2015;10(12):2004-2015. doi:10.1038/nprot.2015.124

10. Köhler S, Carmody L, Vasilevsky N, et al. Expansion of the Human Phenotype Ontology (HPO) knowledge base and resources. *Nucleic Acids Res*. 2019;47(D1):D1018-D1027. doi:10.1093/nar/gky1105

11. Jagadeesh KA, Birgmeier J, Guturu H, et al. Phrank measures phenotype sets similarity to greatly improve Mendelian diagnostic disease prioritization. *Genet Med*. 2019;21(2):464-470. doi:10.1038/s41436-018-0072-y

12. Undiagnosed Diseases Network, Deisseroth CA, Birgmeier J, et al. ClinPhen extracts and prioritizes patient phenotypes directly from medical records to expedite genetic disease diagnosis. *Genet Med*. 2019;21(7):1585-1593. doi:10.1038/s41436-018-0381-1

13. Ewans LJ, Schofield D, Shrestha R, et al. Whole-exome sequencing reanalysis at 12 months boosts diagnosis and is cost-effective when applied early in Mendelian disorders. *Genet Med*. 2018;20(12):1564-1574. doi:10.1038/gim.2018.39

14. Xiao B, Qiu W, Ji X, et al. Marked yield of re-evaluating phenotype and exome/target sequencing data in 33 individuals with intellectual disabilities. *Am J Med Genet A*. 2018;176(1):107-115. doi:10.1002/ajmg.a.38542

15. Birgmeier J, Haeussler M, Deisseroth CA, et al. AMELIE speeds Mendelian diagnosis by matching patient phenotype and genotype to primary literature. *Sci Transl Med*. 2020;12(544):eaau9113. doi:10.1126/scitranslmed.aau9113

16. on behalf of the DDD Study, Wright CF, McRae JF, et al. Making new genetic diagnoses with old data: iterative reanalysis and reporting from genome-wide data in 1,133 families with developmental disorders. *Genet Med*. 2018;20(10):1216-1223. doi:10.1038/gim.2017.246

17. Bergant G, Maver A, Lovrecic L, Čuturilo G, Hodzic A, Peterlin B. Comprehensive use of extended exome analysis improves diagnostic yield in rare disease: a retrospective survey in 1,059 cases. *Genet Med*. 2018;20(3):303-312. doi:10.1038/gim.2017.142