Clinical Evaluation of an AI System for Streamlined Variant Interpretation in Genetic Testing

Jiri Ruzicka[1,#], Jean-Marie Ravel[1,#], Jérôme Audoux[1], Alexandre Boulat[1], Julien Thévenon[2], Kévin Yauy[3], Marine Dancer[4], Laure Raymond[4], Yannis Lombardi[5], Nicolas Philippe[1], Michael GB Blum[1,2], Nicolas Duforet-Frebourg[1,*], Laurent Mesnard[5,6,7]

[1] SeqOne Genomics, Montpellier, France
[2] Institute of Advanced Biosciences, Univ Grenoble Alpes, CNRS UMR 5309, Grenoble, France.
[3] University of Montpellier, LIRMM, CNRS, France
[4] Service de Génétique, Eurofins Biomnis, Lyon, France
[5] Service des Soins intensifs Néphrologiques et Rein Aigu, Hôpital Tenon, Assistance Publique- Hôpitaux de Paris, Paris, France
[6] Centre Maladie rare MAHREA and ERKNET, Hôpital Tenon, Paris, France
[7] INSERM CORAKID, Hopital Tenon, Paris, France

# These authors contributed equally to this work
* Corresponding author

Abstract

The growing use of exome/genome sequencing to diagnose hereditary diseases has increased the interpretive workload for clinical laboratories. Efficient methods are needed to maximize diagnostic yield without overwhelming resources. We developed DiagAI, an AI-powered system trained on 2.5 million ClinVar variants to predict ACMG pathogenicity classes. DiagAI ranks variants, proposes diagnostic shortlists, and identifies probands likely to receive molecular diagnoses. It is powered by UP^2, a machine learning–based score that predicts pathogenicity at the variant level, integrating molecular features, inheritance patterns, and phenotypic data when available. We retrospectively analyzed 966 exomes from a nephrology cohort, including 196 with causal variants and 770 undiagnosed cases. To benchmark UP^2's performance, we evaluated the ranking of 62 causal missense variants. UP^2 ranked variants most effectively beyond shortlist sizes of 10 and identified pathogenic variants missed by AlphaMissense. While REVEL performed well for shortlists of 1–10 variants, it showed lower sensitivity than UP^2 as list size increased. AlphaMissense and CADD were less effective overall. DiagAI identified 94.9% of causal variants in diagnostic exomes with HPO terms, compared to 90.8% without, with median shortlist sizes of 12 and 9 variants, respectively. It achieved a sensitivity of 57.1% and a specificity of 92.6% in tagging exomes likely to contain a diagnostic variant. With HPO terms, 74% of top-ranked variants were diagnostic, versus 42% without, and DiagAI outperformed Exomiser and AI-MARRVEL. DiagAI produces accurate shortlists that streamline variant interpretation and offers a scalable solution for growing diagnostic volumes without compromising quality.

Introduction

Hereditary diseases are a significant health concern worldwide, and exome sequencing (ES) and whole genome sequencing (WGS) have become essential for their diagnosis[1,2]. Efficient genomic variant interpretation is a critical step in clinical genomics[3].

Molecular diagnosis requires interpreting numerous variants detected by high throughput sequencing. The American College of Medical Genetics and Genomics (ACMG) offers standardized guidelines for variant classification. However, these guidelines can lead to discordant classifications among laboratories for a given variant[4] due to the absence of an universal algorithm or precisely defined numerical thresholds[5].

To address these challenges, artificial intelligence (AI) has emerged as a promising solution. Several studies have demonstrated AI's potential to improve the efficiency of variant interpretation workflows, reduce analysis times, and alleviate the human workload[6–9].

We have developed DiagAI, a commercially available AI-powered system designed to predict ACMG variant classifications and prioritize the most likely causal variants. DiagAI relies on a classifier trained on 2.5 million variants annotated in ClinVar[10]. Furthermore, DiagAI upweights genes associated with reported phenotypes to enhance diagnostic precision.

To evaluate the performance of DiagAI, we conducted a retrospective analysis of 966 exomes from adult nephrology patients. DiagAI successfully shortlisted diagnostic variants in 94.9% of cases with HPO-coded phenotypes, compared to 90.8% without. The average shortlist contained 9 variants with HPO terms and 12 without.

Materials and Methods

Study design

To validate DiagAI, we conducted a retrospective analysis from exome sequencing (ES) data generated from 966 adult participants with nephropathy of unknown origin, sequenced between March 2018 and July 2022 (Supplemental Table 1). The cohort included 196 positive cases (24%) with a previously identified causal variant and 770 negative cases (76%) where geneticists could not establish a diagnosis based on ES data.

Exome sequencing

DNA was extracted from peripheral blood using the QIAGEN QIAsymphony DSP DNA Mini Kit. Library preparation and capture was performed with Twist reagents (Human Comprehensive Exome or Human Exome 2.0 Plus Comprehensive Exome Spike-in). Sequencing was performed on the Illumina NovaSeq6000 in paired-end mode (2×150 bp reads). Raw data (bcl format) were converted to FASTQ format using BCL Convert. Reads were aligned to the hg37 human reference genomewith Burrows-Wheeler Aligner Variant calling was performed using SeqOne Genomics' GermVar pipeline.
The GermVar pipeline employs multiple variant callers based on the analysis type. For panel analyses, it uses Freebayes (≥5% allele frequency), GATK, GRIDSS, AluMEI, and GATK-Mutect2-Mitochondrial (≥2.0). Exome analysis utilizes Freebayes (≥10% allele frequency for SNVs), GRIDSS, AluMEI, and GATK-Mutect2-Mitochondrial.

DiagAI priorisation algorithm

SeqOne DiagAI is a tool suite prioritizing genetic variants (SNVs and indels) using molecular and phenotypic data.

A molecular pathogenicity score, UP[2] (Universal Pathogenicity Predictor), is computed using a machine-learning classifier trained on 2.5 million ClinVar variants labeled with ACMG categories. Each variant is represented by 72 molecular features derived from various evidence sources and related to the ACMG criteria (PVS1, PS1, *etc.*). The feature values were calculated using information from several databases including ClinVar for variant annotations and interpretations, Gnomad v4.1 for population frequencies and constraint metrics, dbNSFP dbscSNV for *in silico* predictors, CI-SpliceAI for in silico splice prediction[11], VEP transcript annotations and effect prediction[12], and RepeatMasker for the distance to repetitive genomic regions (related to the PM4 criterion).

A comprehensive prioritization score, referred to as the DiagAI score, was built upon the UP[2] molecular score by incorporating additional contextual information when available. This includes phenotypic data, variant inheritance mode consistency from databases such as PanelApp[13], OMIM, and MedGen, variant calling data (DP, AO, VAF) and quality (base quality phred score), and parental variant data in cases where trio sequencing was performed. Phenotypic information was incorporated using HPO terms, with gene prioritization performed by Phenogenius, a tool that leverages gene-HPO term associations from literature-based matrices[14]. A machine learning regression model, trained on 307

diagnosed exomes (26 million variants), integrates ACMG-related molecular features with HPO-based phenotypic features.

A shortlist of likely causal variants is generated by applying thresholds to both the molecular UP^2 score and the comprehensive DiagAI score. Variants with a frequency above 1.5% in the phenotypically matched cohort are discarded.

DiagAI also tags exomes/genomes likely to yield a molecular diagnosis. An exome is considered tagged if it contains at least one 'smartpick' variant for which both the DiagAI and UP^2 scores exceed predefined thresholds that are more stringent than those used to define the shortlist. While a shortlist is generated for all exomes/genomes, only some are tagged as likely to lead to a diagnosis.

Interpretability features

To assess the importance of different ACMG criteria in our classification model, we calculated Shapley values for all 72 features used in the UP^2 molecular pathogenicity score. These Shapley values were then aggregated and mapped to the known ACMG criteria. This approach allowed us to quantify each criterion's contribution to the molecular UP^2 score and indicate their influence on pathogenicity prediction within our model. The Shapley value for each ACMG criterion was calculated by summing the Shapley values of its associated features.

Evaluation of performance

We compared the UP^2 score's ability to rank missense diagnostic variants against CADD, REVEL and Alphamissense[15–17]. We focused on missense variants, for which predictive algorithms are key to ACMG classification. To ensure an unbiased evaluation, diagnostic variants present in ClinVar (used as training data for UP^2) were excluded.
To benchmark the performance of DiagAI, we compared its ranking performance to AI-MARRVEL (lite version, no HGMD) and Exomiser (v13 with the following parameters: MVP and REVEL pathogenicity sources, failedVariantFilter, inheritanceFilter, frequencyFilter and pathogenicityFilter with the keepNonPathogenic option set to true) [6,18].
Since DiagAI uses quality, variant allele frequency, depth and number of observed alternate alleles, we applied identical filtering prior to the run of Exomiser13 and AI-MARRVEL. We evaluated rankings at the gene level. For DiagAI and AI-MARRVEL, genes were ranked based on their highest-priority variant, whereas Exomiser already provides a ranked list of genes.

Results

Comparison between UP^2, CADD, REVEL AND AlphaMissense

We compared the rankings of 62 causal missense variants (not present in ClinVar) by UP^2 AlphaMissense, CADD and REVEL (Figure 1). For shortlists of ≤10 variants, REVEL showed the best prioritization, ranking the causal variant first in 12.90% of cases and within the top 10 in 32.26%. UP^2 achieved 0% top-1 and 27.42% top-10 inclusion, while both AlphaMissense and CADD showed 0% top-1 and 1.61% top-10 inclusion. For larger lists (>10 variants), UP^2 outperformed other methods, including the causal variant within the top 100 in 87.10% of cases, compared to 61.29% for REVEL, 19.35% for AlphaMissense, and 11.29% for CADD (Figure 1).

Direct comparison of AlphaMissense and UP^2 scores revealed 72.6% (45/62) concordance in pathogenic classification. However, AlphaMissense incorrectly classified 11.3% (7/62) of causal variants as benign, which UP2 correctly identified as pathogenic (Figure 1, lower left panel). No variants were considered as pathogenic by Alphamissense and benign with UP^2. REVEL and UP^2 also showed strong concordance, with only one causal variant classified as benign by REVEL and as VUS by UP^2, and no pathogenic/benign discrepancies (Figure 1, lower right panel).

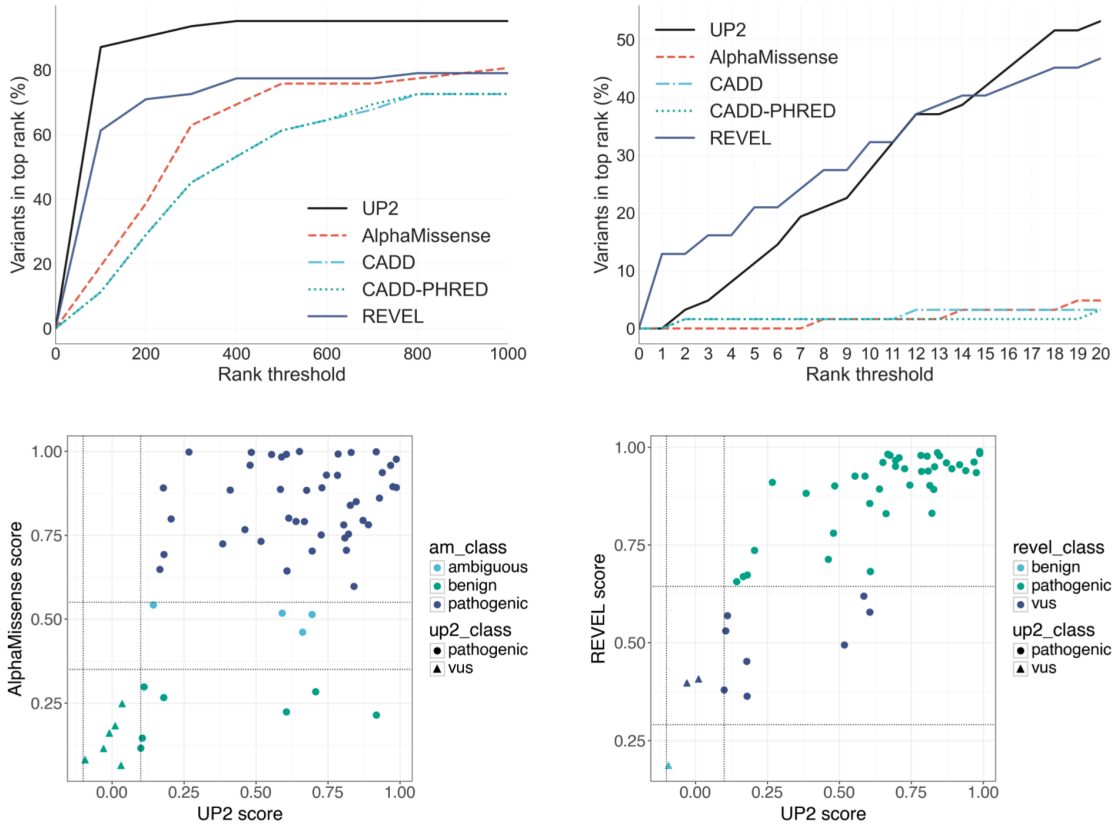


Figure 1: Performance of UP^2 variant ranking evaluated on 62 missense variants not reported in ClinVar. Top-panels : percentage of analysis with the diagnostic variant being in the top-rank list of variants (left: top 1000, right: top 20) according to different scoring systems: AlphaMissense, CADD, REVEL and UP^2. Bottom-level panels : Variants' scores comparison between UP^2 and AlphaMissense (left) or REVEL (right) (62 causal missense variants not in ClinVar). am_class : classification predicted by AlphaMissense ; up2_class : classification predicted by UP^2 ; am_revel: classification predicted by REVEL. Dashlines indicate the threshold of corresponding algorithms. The down right corner or Alphamissense vs UP^2 comparison corresponds to the variants considered benign by AlphaMissense but reported as pathogenic by UP^2.

Proportion of Causal Variants Identified in Shortlists

In exomes with a confirmed diagnosis, the causal variant was present in the DiagAI shortlist in 94.9% (186/196) of cases when HPO terms were included, versus 90.8% (176/196) without HPO terms. The median shortlist size was 9 variants (min=2, max=44) without HPO terms and 12 variants (min=4, max=29) with HPO terms.

Sensitivity and Specificity of DiagAI in Diagnosing Exomes

DiagAI tagged 57.1% (112/196) of probands with a molecular diagnosis as likely diagnosable (sensitivity). Among the 770 undiagnosed exomes, DiagAI correctly excluded 713 and incorrectly tagged 57, yielding a specificity of 92.6%.

Variant ranking

We assessed DiagAI's variant ranking accuracy with and without HPO-based clinical data on 196 diagnosed exomes (Figure 2). Incorporating clinical data significantly improved the ranking accuracy for the top-ranked variant increasing from 42% diagnostic without HPO terms to 74% with HPO terms The improvement was smaller for the top 20 genes (93% without HPO vs. 97% with HPO terms). DiagAI showed better ranking performance than Exomiser and AI-MARRVEL when HPO terms were used, and for top-ranked lists of three or more variants when HPO terms were not included (Figure 2).

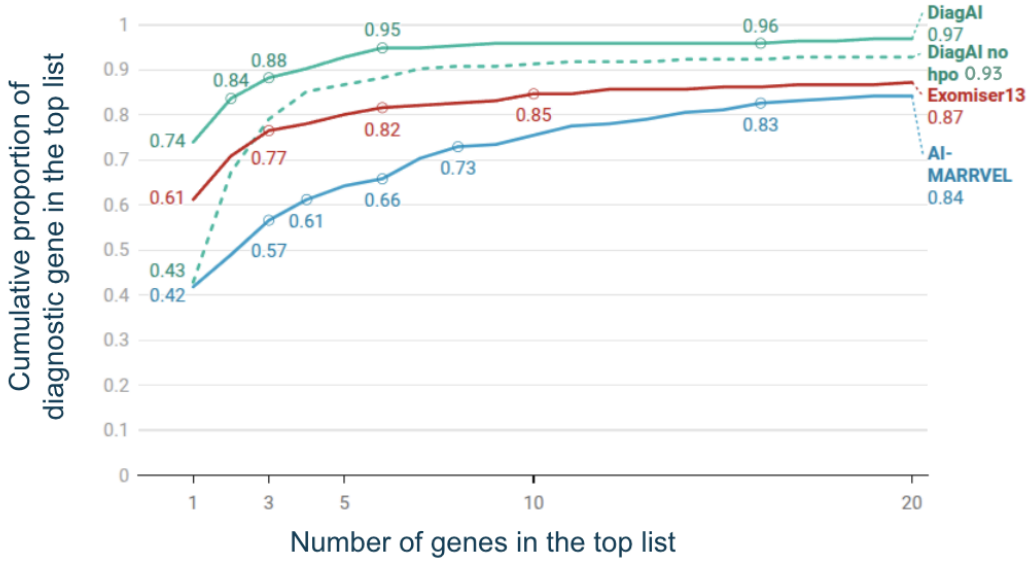


Figure 2. Performance of variant ranking evaluated on the 196 exomes with a confirmed molecular diagnosis. The ranking accuracy of top-ranked genes was assessed using two approaches when evaluating DiagAI: the molecular pathogenicity score alone (no HPO) and a comprehensive score integrating molecular pathogenicity and HPO-based clinical information (with HPO). Variants with a cohort frequency above 1.5% were excluded from the DiagAI ranking. Exomiser and AI-MARRVEL, which also use HPO for gene ranking, were used for comparison.

Interpretability of the classifier

We used the interpretability framework to identify the most influential ACMG criteria for 176 diagnostic variants from 196 diagnosed exomes (several diagnostic variants were identical).

For 85 diagnostic variants with ClinVar submissions, features related to the number of pathogenic and benign submissions were the most influential (Figure 3). However, for 13 variants, other ACMG criteria were more impactful, including 6 with PP3/BP4 (in-silico predictors of pathogenicity and benignity), 5 with PVS1 (predicted impact by VEP), and 2 with PM2/BA1 (absent or at extremely low frequency in general population).

For the 91 diagnostic variants without ClinVar submissions, PP3/BP4 features were instrumental for 66% (60/91) of the variants, while PVS1 features were key for 29% (26/91). The remaining five variants' scores were mainly determined by a combination of other ACMG features.

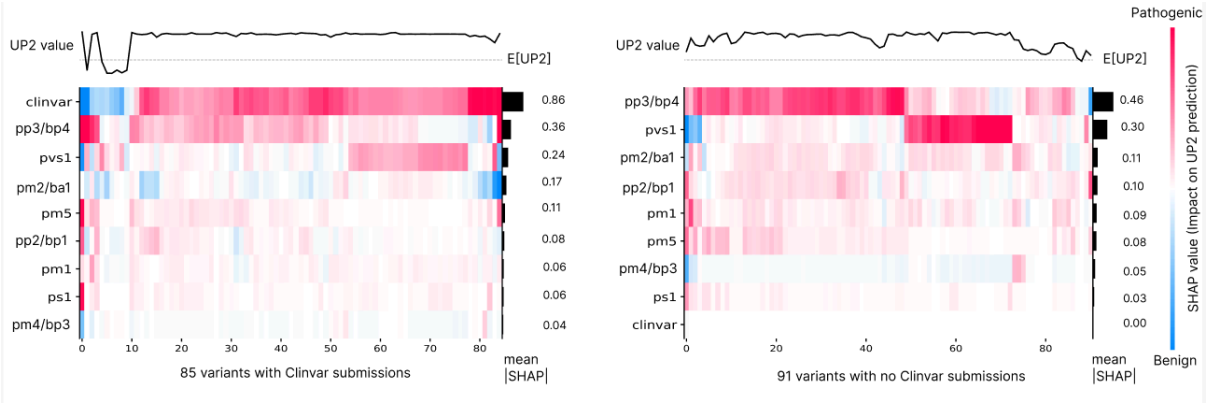


Figure 3. Explicability profile of the 176 diagnostic variants for the Universal Pathogenicity Predictor. The heatmap shows for each variant if the impact of the features related to an ACMG criteria is rather pathogenic (red) or benign (blue). Variants that have ClinVar Submissions (Left panel) have UP^2 predictions that are mostly influenced by the ClinVar submission predictors with an average absolute SHaP value of .86 for ClinVar predictors. Variants that have no ClinVar Submission (Right panel) have UP^2 predictions that are mostly influenced by features from PP3/BP4 or PVS1 criteria. The right side histogram shows that the average absolute SHaP value is .46 for PP3/BP4 and .30 for PVS1 on the UP^2 value for variants without ClinVar submission. The top plots show UP^2 values that are between -1 and 1.

Causal variants absent from the shortlists

Ten diagnostic variants identified by geneticists were not included in the DiagAI shortlists. Five of these were classified as benign in ClinVar, and three were VUS or had conflicting ClinVar classifications without further pathogenicity evidence. The remaining two were missed due to a low-quality variant and a recessive gene variant lacking a sufficiently high-scoring second heterozygous variant to support a composite diagnosis.

Discussion

Our study demonstrates DiagAI's effectiveness in prioritizing causal variants in single-case exome analyses of nephrology patients, with 90-96% of shortlists containing the causal variant depending on phenotypic data availability. DiagAI also outperformed Exomiser and AI-MARRVEL in gene-level ranking.

Comparing UP^2 to other variant scoring tools for causal missense variants not present in ClinVar, REVEL excelled in top variant ranking but its sensitivity decreased with larger shortlists. In contrast, UP2 maintained strong performance across broader ranking ranges, which is crucial for maximizing diagnostic sensitivity. AlphaMissense and CADD showed consistently lower performance, underscoring the added value of UP^2—and to a lesser extent REVEL—value in variant prioritization. While tools like AlphaMissense, CADD, REVEL, and UP^2 focus solely on predicting the pathogenicity of individual variants, diagnostic software such as Exomiser, AI-MARVELL, and DiagAI integrate these scores with patient phenotype and inheritance patterns, offering more clinically relevant prioritization.

DiagAI outperforming open source alternatives is explained by several key aspects. First, the UP^2 algorithm was meticulously designed by selecting and analyzing a comprehensive set of features to enhance its ability to identify pathogenic variants. Rigorous curation of ClinVar data, including the resolution of conflicting classifications, also contributed significantly[19]. Furthermore, we apply data augmentation to mitigate biases, particularly the overrepresentation of pathogenic variants in ClinVar, which can distort model performance, especially in intergenic regions, where benign variants are underreported. Secondly, DiagAI combines machine learning models trained on patient cohorts with expert-driven methods to capture complex diagnostic patterns. This hybrid approach addresses key challenges, such as accurately assessing variant inheritance — especially in compound heterozygous cases — evaluating data quality, and managing differences across sequencing technologies. By integrating these elements, DiagAI enhances diagnostic reliability and remains adaptable to various data types, including long-read sequencing.

We examined some of the causal variants missed by the shortlist of DiagAI. We identified one variant in the PODXL gene, whose pathogenic effect has been linked to kidney diseases[20]. We also identified a *CFI* variant, dysregulating the complement alternative pathway. However, *CFI* variants are challenging to interpre and to classify using ACMG criteria[21], often considered risk factors rather than Mendelian causes[22]. Additionally, a

synonymous variant in *NPHP3* with a splicing effect that is not captured by computational approaches, was also overlooked[23]. Overall, DiagAI missed these variants, which geneticists diagnosed using recent research and expert knowledge of kidney disease genes.

These cases highlight the inherent challenges in interpreting variants with emerging or complex pathogenic mechanisms not yet fully captured by computational models. While DiagAI improves variant prioritization by leveraging multivariate ACMG evidence tags and machine learning trained on ClinVar ACMG classifications, it remains limited by current knowledge and training data. Integrating diverse evidence and HPO-coded phenotypes enhances ranking, but expert review remains essential for novel or challenging cases.

DiagAI joins the expanding field of of computational solutions for variant prioritization, joining both open-source and commercial offerings such as AI-MARRVEL, InVitae MOON, Fabric GEM, and the Emedgne software from Illumina[6,8,9,24]. Direct comparisons between these tools are challenging due to differences in their evaluation cohorts. Nonetheless, we found that DiagAI outperformed AI-MARRVEL in gene ranking and demonstrated comparable performance in identifying diagnosable cases, with both tools automatically detecting 50–60% of such cases[6]. While the CAGI challenge provides a standardized benchmarking framework, DiagAI has not yet been evaluated through it[25].

DiagAI's accuracy in variant ranking, particularly when integrating clinical data, highlights its potential to streamline genomic diagnostics by reducing the number of variants requiring manual review. However, the path to full automation remains long, with less than 60% of diagnosed cases detected automatically. These findings suggest that AI-powered tools like DiagAI can significantly reduce the interpretive workload in clinical genomics while maintaining high diagnostic accuracy. This assessment should be extended beyond the specific case of nephrology and tested in whole genome sequencing analysis.

Bilbiography

1. 100,000 Genomes Pilot on Rare-Disease Diagnosis in Health Care — Preliminary Report. *N Engl J Med*. 2021;385(20):1868-1880. doi:10.1056/NEJMoa2035790

2. Doreille A, Lombardi Y, Dancer M, et al. Exome-First Strategy in Adult Patients With CKD: A Cohort Study. *Kidney Int Rep*. 2023;8(3):596-605. doi:10.1016/j.ekir.2022.12.007

3. Austin-Tse CA, Jobanputra V, Perry DL, et al. Best practices for the interpretation and reporting of clinical whole genome sequencing. *NPJ Genomic Med*. 2022;7(1):27.

4. Amendola LM, Muenzen K, Biesecker LG, et al. Variant Classification Concordance

using the ACMG-AMP Variant Interpretation Guidelines across Nine Genomic Implementation Research Studies. *Am J Hum Genet*. 2020;107(5):932-941. doi:10.1016/j.ajhg.2020.09.011

5. Pejaver V, Byrne AB, Feng BJ, et al. Calibration of computational tools for missense variant pathogenicity classification and ClinGen recommendations for PP3/BP4 criteria. *Am J Hum Genet*. 2022;109(12):2163-2177.

6. Mao D, Liu C, Wang L, et al. AI-MARRVEL — A Knowledge-Driven AI System for Diagnosing Mendelian Disorders. *NEJM AI*. 2024;1(5). doi:10.1056/AIoa2300009

7. Owen MJ, Lefebvre S, Hansen C, et al. An automated 13.5 hour system for scalable diagnosis and acute management guidance for genetic diseases. *Nat Commun*. 2022;13(1):4057.

8. De La Vega FM, Chowdhury S, Moore B, et al. Artificial intelligence enables comprehensive genome interpretation and nomination of candidate diagnoses for rare genetic diseases. *Genome Med*. 2021;13(1):153. doi:10.1186/s13073-021-00965-0

9. Meng L, Attali R, Talmy T, et al. Evaluation of an automated genome interpretation model for rare disease routinely used in a clinical genetic laboratory. *Genet Med*. 2023;25(6):100830.

10. Landrum MJ, Chitipiralla S, Brown GR, et al. ClinVar: improvements to accessing data. *Nucleic Acids Res*. 2020;48(D1):D835-D844. doi:10.1093/nar/gkz972

11. Strauch Y, Lord J, Niranjan M, Baralle D. CI-SpliceAI—Improving machine learning predictions of disease causing splicing variants using curated alternative splice sites. *PLOS ONE*. 2022;17(6):e0269159. doi:10.1371/journal.pone.0269159

12. McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. *Genome Biol*. 2016;17(1):122. doi:10.1186/s13059-016-0974-4

13. Martin AR, Williams E, Foulger RE, et al. PanelApp crowdsources expert knowledge to establish consensus diagnostic gene panels. *Nat Genet*. 2019;51(11):1560-1565. doi:10.1038/s41588-019-0528-2

14. Yauy K, Duforet-Frebourg N, Testard Q, et al. Learning phenotypic patterns in genetic disease by symptom interaction modeling. Published online July 31, 2022:2022.07.29.22278181. doi:10.1101/2022.07.29.22278181

15. Cheng J, Novati G, Pan J, et al. Accurate proteome-wide missense variant effect prediction with AlphaMissense. *Science*. 2023;381(6664):eadg7492. doi:10.1126/science.adg7492

16. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res*. 2019;47(D1):D886-D894. doi:10.1093/nar/gky1016

17. Ioannidis NM, Rothstein JH, Pejaver V, et al. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *Am J Hum Genet*. 2016;99(4):877-885. doi:10.1016/j.ajhg.2016.08.016

18. Cipriani V, Pontikos N, Arno G, et al. An Improved Phenotype-Driven Tool for Rare Mendelian Variant Prioritization: Benchmarking Exomiser on Real Patient Whole-Exome

Data. *Genes*. 2020;11(4):460. doi:10.3390/genes11040460

19. Yauy K, Lecoquierre F, Baert-Desurmont S, et al. Genome Alert!: A standardized procedure for genomic variant reinterpretation and automated gene-phenotype reassessment in clinical routine. *Genet Med Off J Am Coll Med Genet*. 2022;24(6):1316-1327. doi:10.1016/j.gim.2022.02.008

20. Blasco M, Quiroga B, García-Aznar JM, et al. Genetic Characterization of Kidney Failure of Unknown Etiology in Spain: Findings From the GENSEN Study. *Am J Kidney Dis Off J Natl Kidney Found*. 2024;84(6):719-730.e1. doi:10.1053/j.ajkd.2024.04.021

21. Schwotzer N, Fakhouri F, Martins PV, et al. Hot Spot of Complement Factor I Rare Variant p.Ile357Met in Patients With Hemolytic Uremic Syndrome. *Am J Kidney Dis*. 2024;84(2):244-249. doi:10.1053/j.ajkd.2023.12.021

22. Timmermans SAMEG, van Doorn DPC, van Paassen P. Rare Variants in Complement Genes May Not Be That Rare After All. *Kidney Int Rep*. 2023;8(10):1911-1913. doi:10.1016/j.ekir.2023.08.020

23. Molinari E, Decker E, Mabillard H, et al. Human urine-derived renal epithelial cells provide insights into kidney-specific alternate splicing variants. *Eur J Hum Genet EJHG*. 2018;26(12):1791-1796. doi:10.1038/s41431-018-0212-5

24. O'Brien TD, Campbell NE, Potter AB, Letaw JH, Kulkarni A, Richards CS. Artificial intelligence (AI)-assisted exome reanalysis greatly aids in the identification of new positive cases and reduces analysis time in a clinical diagnostic laboratory. *Genet Med*. 2022;24(1):192-200.

25. Critical Assessment of Genome Interpretation Consortium. CAGI, the Critical Assessment of Genome Interpretation, establishes progress and prospects for computational genetic variant interpretation methods. *Genome Biol*. 2024;25(1):53. doi:10.1186/s13059-023-03113-6

Data Availability

Data should be requested to Prof L Mesnard.

Conflict of Interest

J. Ruzicka, J.-M. Ravel, J. Audoux, A. Boulat, K. Yauy, N. Philippe, M. Blum, and N. Duforet-Frebourg are current or former employees of SeqOne Genomics and hold, or have received, stock or stock options from the company. M. Dancer and L. Raymond are employees of Eurofins Biomnis.

Acknowledgments

We would like to thank the technicians, scientists, bioinformaticians, operations staff, and biologists at Eurofins Biomnis for their valuable contributions to the production and interpretation of these data.

Funding

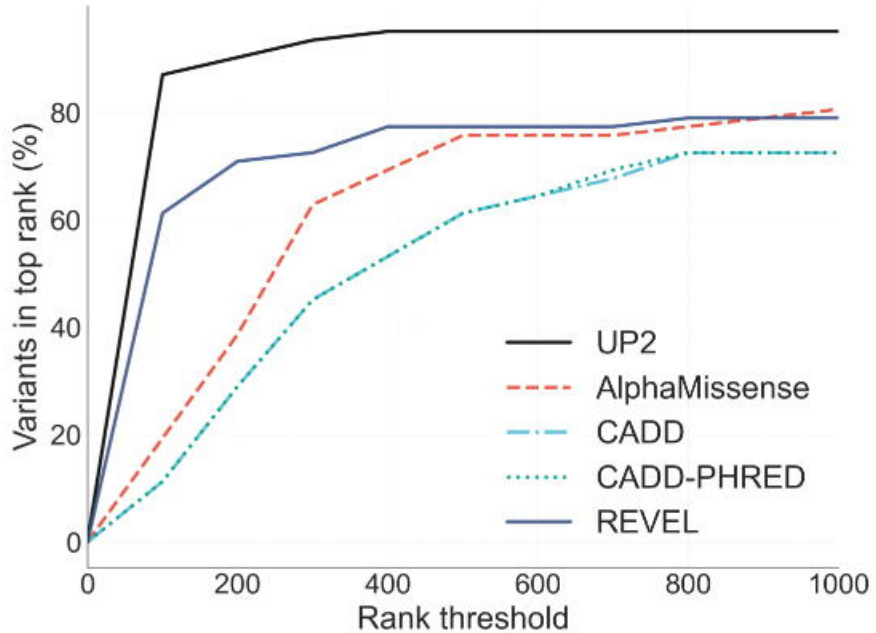
This study was conducted without additional funding.

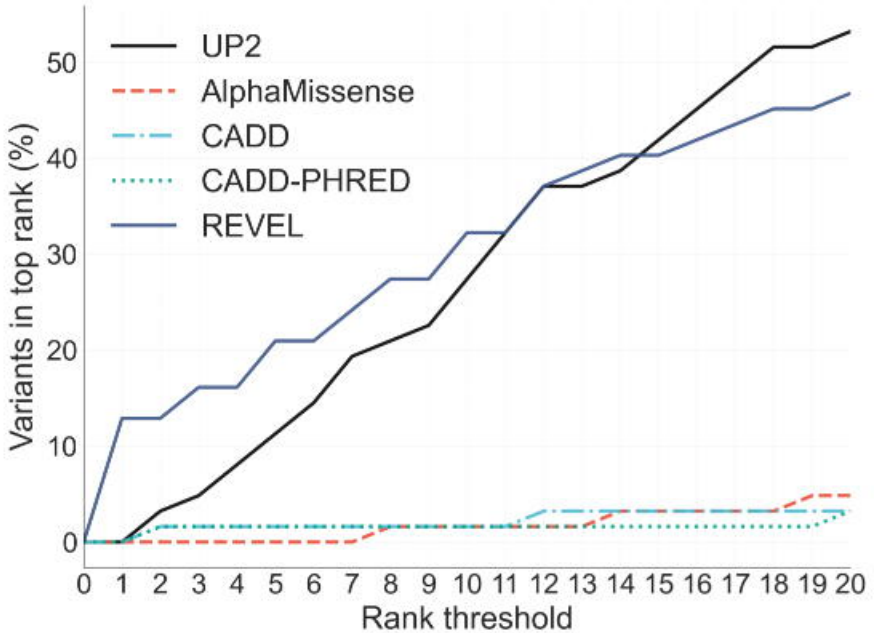
Author Information

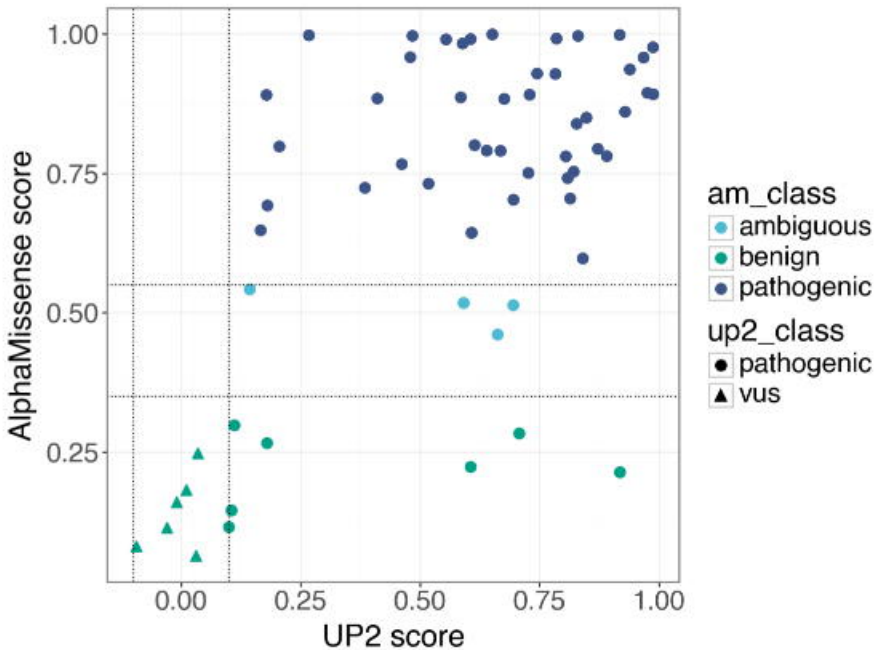
Conceptualization: M.G.B., N.P., N.D.-F., L.M.; Methodology: J.R., K.Y., N.D.-F., J.T.; Data Curation: J.R., A.B.; Data Analysis: J.R., N.D.-F., J.-M.R.; Data Production: L.M., M.D., Y.L., L.R.; Software: J.R., N.D.-F., J.A., N.P.; Supervision: M.G.B., N. D.-F.; Writing – Original Draft: M.G.B.; Writing – Review and Editing: M.G.B., J.-M.R., L.M.

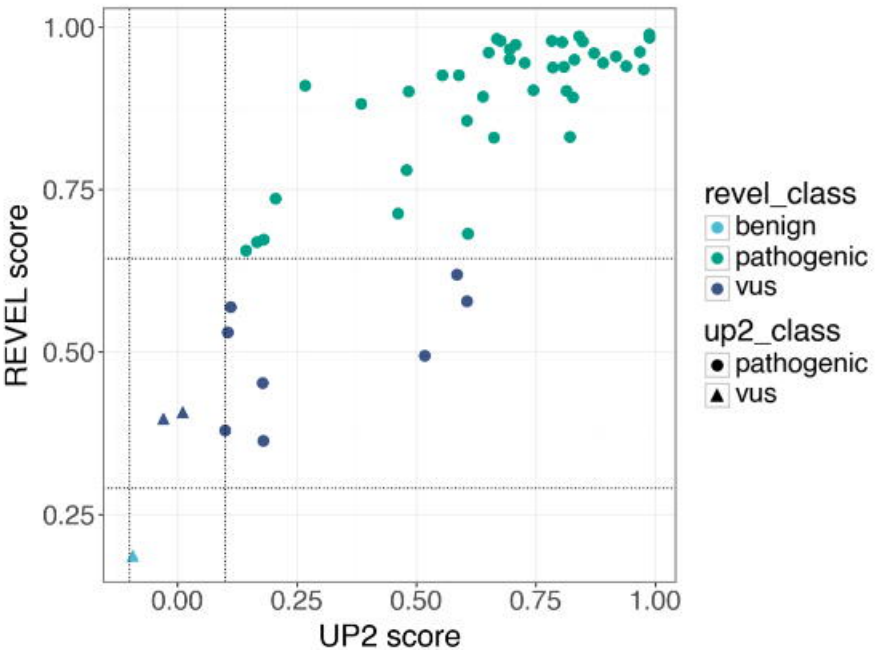
Ethics Statement

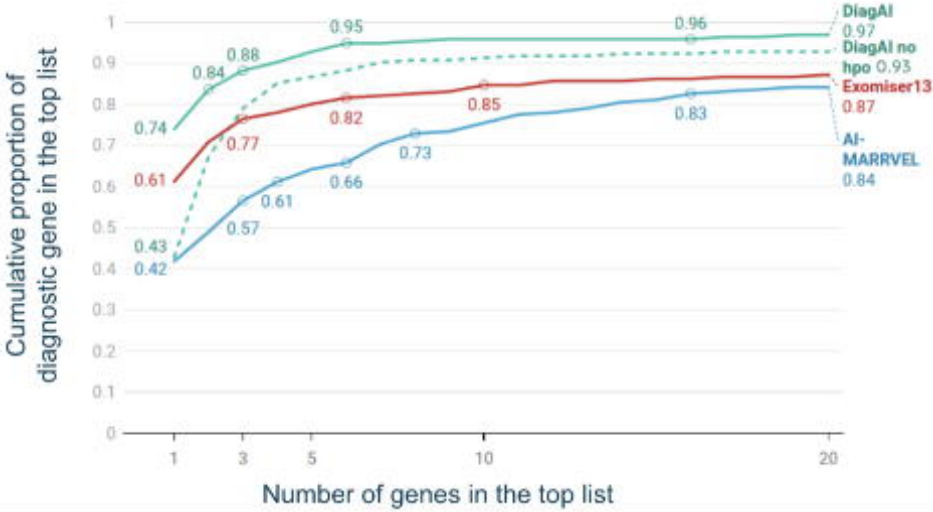
The cohort has been previously described[2]. The study was approved by an Institutional Review Board (Direction de la Recherche Clinique et de l'Innovation (APHP220461) and the Ethic board of Sorbonne Université (CER-2022-009).

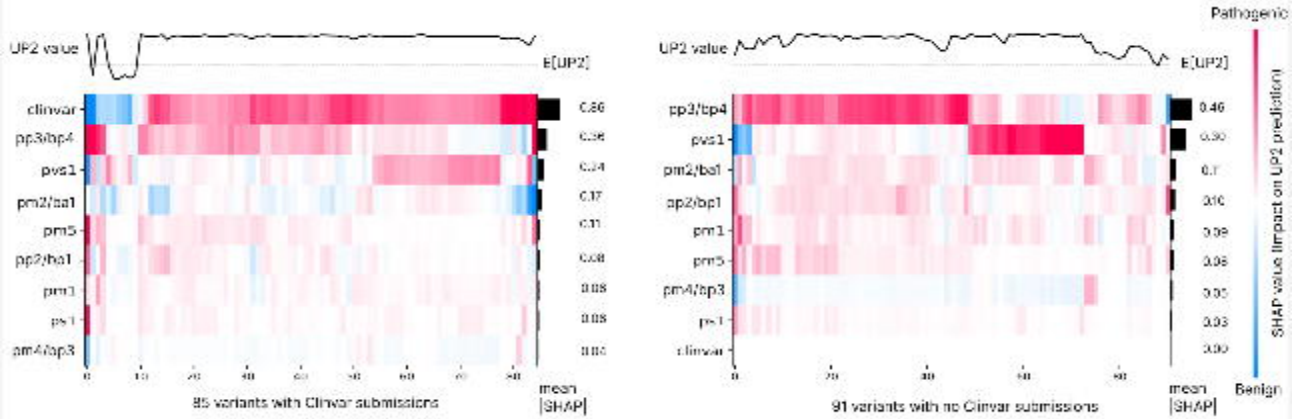

Variants in top rank (%)
Rank threshold
UP2
AlphaMissense
CADD
CADD-PHRED
REVEL


Variants in top rank (%)
Rank threshold
UP2
AlphaMissense
CADD
CADD-PHRED
REVEL


AlphaMissense score
UP2 score
am_class
ambiguous
benign
pathogenic
up2_class
pathogenic
vus


REVEL score
UP2 score
revel_class
benign
pathogenic
vus
up2_class
pathogenic
vus


Cumulative proportion of diagnostic gene in the top list
Number of genes in the top list
DiagAI
DiagAI no hpo
Exomiser13
AI-MARRVEL
0.74
0.84
0.88
0.95
0.96
0.97
0.93
0.87
0.84
0.61
0.77
0.82
0.85
0.83
0.43
0.42
0.57
0.61
0.66
0.73


UP2 value
E[UP2]
clinvar
pp3/bp4
pvs1
pm2/ba1
pm5
pp2/ba1
pm1
ps1
pm4/bp3
mean
|SHAP|
85 variants with Clinvar submissions
UP2 value
E[UP2]
pp3/bp4
pvs1
pm2/ba1
pp2/bp1
pm1
pm5
pm4/bp3
ps1
clinvar
mean
|SHAP|
91 variants with no Clinvar submissions
Pathogenic
SHAP value (impact on UP2 prediction)
Benign