Nostos Genomics

Reducing complexity in variant interpretation through AI

A CLINICAL VALIDATION STUDY OF AION ON THE GENOMICS ENGLAND 100.000 GENOMES PROJECT

www.nostos-genomics.com

Nostos Genomics

Turn data into genomic insight

Table of contents

In this white paper:
How new developments in machine learning support scalability in genomic research.

Genome sequencing has the potential to resolve rare undiagnosed diseases. Yet, modern genomic tests do not lead to a diagnosis for every patient. Many patients undergo a diagnostic odyssey prolonged by the challenge of reviewing hundreds of DNA variants to assess their implications for disease. To overcome the variant interpretation bottleneck, clinical laboratories need rapid technological innovation. The next frontier in genomic testing lies in solutions that leverage artificial intelligence (AI) to improve existing processes. AION is the flagship AI-driven platform for rare disease variant interpretation, developed by Nostos Genomics. AION supports analysts to diagnose rare diseases faster using a machine-learning model trained on millions of high-quality genetic variant data points. When applied to cases from the 100,000 Genomes Project, AION identified the causative variant in 91.5% of cases, increasing to 93.1% when parental information is provided and 94% in paediatric patients. AION empowers clinical laboratories and genetic testing services to deliver accurate, timely, reproducible variant interpretations for rare disease patients.

In this white paper, we uncover answers to key questions in rare disease genomics:

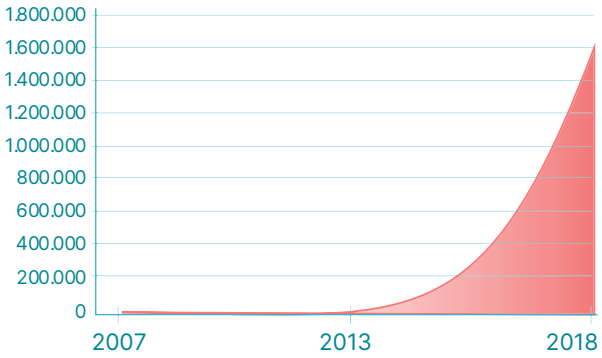
- What are the current barriers to accurate variant interpretation?
- What is the role of artificial intelligence in scaling variant interpretation?
- How does AION modernise genomic testing for rare diseases?

Genomic testing in rare disease:
Slow, complex and costly.

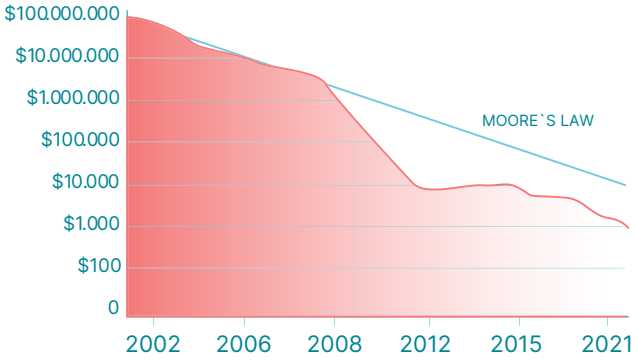
Three advances in human genetics research have enabled us to link genetic variants to rare diseases: the completion of the reference genome in 2003, providing a map from which we can find genetic differences [1]; the advent of next-generation sequencing (NGS) technologies which enabled high-throughput profiling of patient DNA at <$1000 per genome [2]; and the development of databases and algorithms for detecting, annotating and interpreting genomic variation [3]. As sequencing costs have decreased, rare disease genetic testing has migrated from small gene panels to whole exome sequencing (WES) for protein-coding regions and whole genome sequencing (WGS) for a patient's complete DNA sequence profile.

Yet, despite these powerful technologies, many rare disease patients do not receive a diagnosis [4,5]. For example, the Genomics England 100,000 Genomes Project (100kGP) pilot study reports that only 25% of 2183 patients received a definitive diagnosis after whole genome sequencing [4]. One driver of this low diagnostic yield is the lack of evidence linking variants to rare disease phenotypes. Many variants lack sufficient evidence for benign or pathogenic classification and are therefore relegated to the ambiguous category of Variants of Uncertain Significance (VUS).

EXPONENTIAL GROWTH IN GENOME SEQUENCING

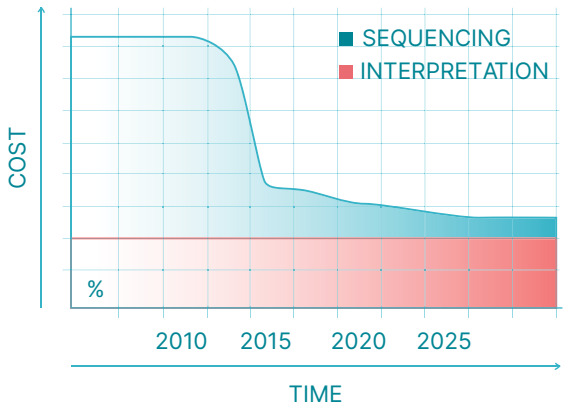


COST PER HUMAN GENOME



A further challenge to diagnosing rare diseases is the variant interpretation bottleneck. Variant interpretation is a manual, time-consuming process. As a result, decision support software has become commonplace in rare disease testing [6]. These applications streamline variant assessment by annotating each variant with information from the broad range of databases consulted by experts, and by supporting the use of ACMG/AMP criteria. To apply ACMG/AMP criteria, experts conduct an extensive search of scientific literature and clinical variant databases. However, WGS returns hundreds to thousands of variants from each patient, a volume that overwhelms personnel resources in genomic testing laboratories. To address these issues, laboratories rely on decision support tools to streamline analysis by ranking variants and automating ACMG/AMP classifications. As the success of rare disease genetic testing lies in overcoming the bottleneck in variant interpretation, this grand challenge in human genetics demands a new class of computational tools that propel experts beyond existing bottlenecks.

INTERPRETATION BECOMES COST DRIVER AS SEQUENCING COSTS DROP



Nostos Genomics

Reducing complexity in variant interpretation through AI.

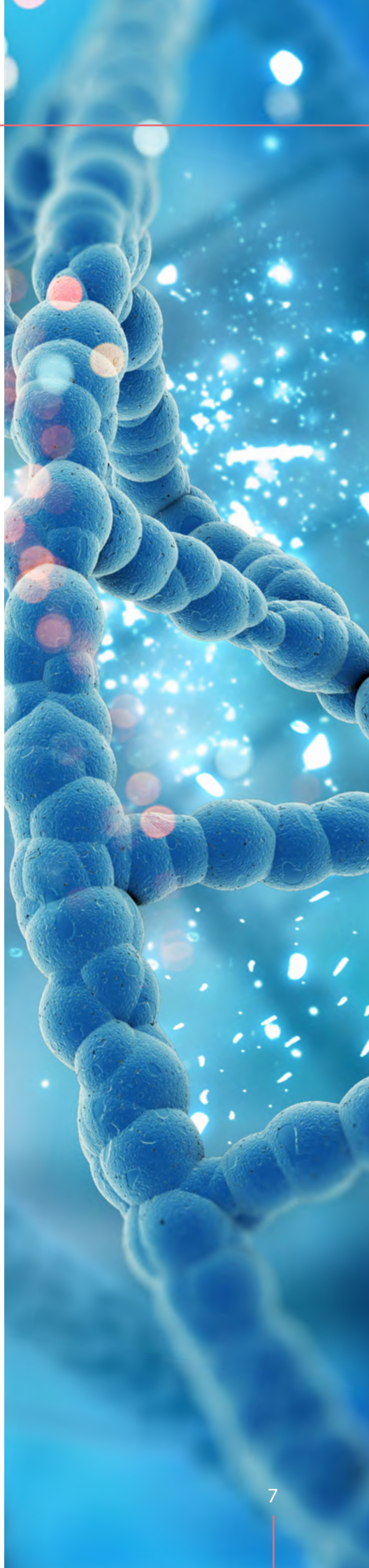
AI-driven variant prioritisation offers three primary benefits to clinical laboratories. First, AI software affords a higher sensitivity for diagnosing rare diseases. As AI models prioritise variants using patterns learned from millions of data points, they are more likely than simple rule-based algorithms to present analysts with truly pathogenic variants. Trained AI models may also be applied to previously unresolved cases, guiding diagnoses that were previously missed. Second, AI software can reduce the time taken to identify a diagnosis. By quickly identifying the most promising candidate variants for interpretation, AI-driven decision support tools lead experts to make faster diagnoses. Lastly, AI places clinical laboratories in a position to scale their operations. The cost of genome sequencing has steadily declined since the $150 million draft human genome in 2003, as evidenced by the $200 genome made possible in 2022 by Illumina's NovaSeq X [7]. Laboratories wishing to scale their operations in light of decreasing sequencing costs will require more analysts. Still, individuals with expertise in genomics may not be in supply to meet the increase in demand. Therefore, AI-driven decision support tools are a workforce multiplier, granting fast and accurate diagnoses that enable rare disease laboratories to serve ever-larger populations. The potential to scale operations would further see the launch of more population-scale genome sequencing and a greater business need for AI tools.

AI algorithms will enable laboratories to push the boundaries of their diagnostic practice by granting analysts more time to focus on complex cases. Although most rare diseases will be solved quickly with AI-driven interpretation, a small number of patients may not show a clear causative variant. These complex cases may present several VUS or highly atypical phenotypes. Clinical laboratories typically perform in-depth investigations into these VUS using functional screens or additional familial sequencing, both of which may provide evidence for reclassification. By prioritising VUS with a pathogenicity score, AI tools add nuance to the VUS category and indicate the most valuable variants for in-depth investigations, allowing genomics departments to make the best use of limited resources.

The AION platform.

At Nostos Genomics, we recognise the need for new solutions that leverage artificial intelligence to end the rare disease diagnostic odyssey. Our research and development team invested decades of expertise in clinical genomics, software engineering and data science to develop AION, an AI-driven decision support platform for the molecular diagnosis of rare diseases. AION takes advantage of emerging technologies available for genomic analysis, bridging the gap between current practice and the promise of personalised medicine towards faster, more accurate diagnoses for rare disease patients.

AION is an all-in-one CE-IVD certified system for variant annotation, variant classification and variant prioritisation. Given variants from gene panels, exomes or whole genomes, AION grants users a streamlined summary of variant annotations. Analysts will find easy access to essential sources, such as Ensembl and Refseq annotations, gene dosage sensitivity scores (from ClinGen), known genetic disease associations, population allele frequencies, predicted molecular consequence, and evolutionary constraint scores. Annotated variants are processed by an implementation of the ACMG/AMP guidelines. In parallel, AION scores genetic variants using our innovative machine-learning classifier. Given details of the patient's disease presentation, AION is capable of mapping these terms to the Human Phenotype Ontology before ranking the most likely genetic diseases associated with the patient's variants and clinical phenotype.



The AION classifier is underpinned by an AI model that is constantly updated. A pathogenicity score is provided for each variant, complete with a breakdown of how the AION model arrives at the score, enabling analysts to collate evidence towards ACMG classifications with ease. At Nostos Genomics, we enhanced AION's training dataset with results from our high-throughput functional genomics screens, allowing AION to intelligently learn from both experimental data and large-scale clinical databases. AION is iteratively refined on a rich repository of ground-truth associations between complex variant features and disease states. When variants cannot be confidently classified, the AION score complements ACMG/AMP guidelines by prioritising VUS for follow-up investigations. Here, AION quickly presents the variants most likely to bring a diagnosis through further investigation.

As genomics laboratories may have limited resources to integrate new software, AION is designed to integrate seamlessly into existing workflows. As a web platform, AION grants users access to its comprehensive features directly from an internet browser, eliminating the need for technical expertise. AION's user-centred design brings analysts high-quality note-taking tools to record the details of each classification. For laboratories looking for deeper integration, AION's application programming interface lets users execute our powerful AI classifier as a new feature within another decision support system directly through API. As clinical decision support software must meet strict governance requirements expected of a medical device, we offer AION with CE-IVD certification, giving confidence in its compliance, stability, and safety for use in diagnostic settings. AION Premium is our more advanced version of the platform, with a constantly updated and continuously improving set of algorithms and tools.

AION leverages AI to provide a feature-rich platform that propels clinical laboratories beyond the interpretation bottleneck towards timely diagnosis, reduced operational costs, and higher sample throughput.

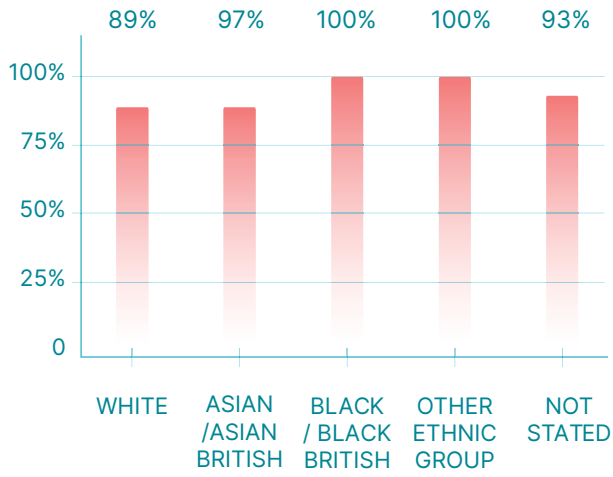
Perfomance of AION on the Genomics England 100.000 Genomes Project.

To validate AION's clinical performance, we analysed patients enrolled on the 100,000 Genomes Project (100kGP), a milestone in clinical genomics whereby 100,000 samples from patients in the United Kingdom's National Health Service (NHS) were sequenced by WGS towards a national genomic medicine service [4].

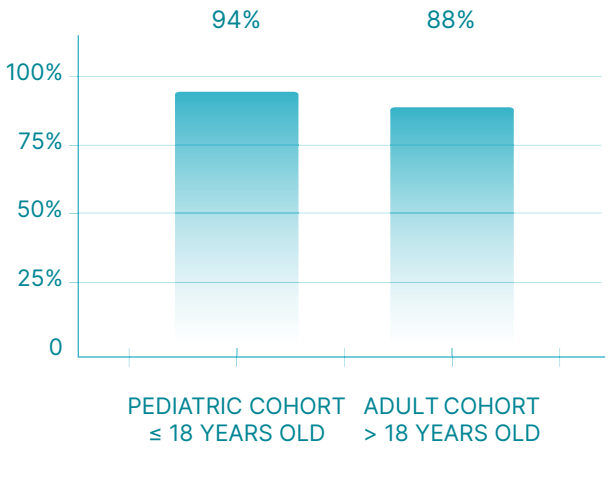
This benchmarking study was performed on consented and solved cases using AION version 3.0. Individuals who underwent WGS analysis with reference genome GRCh37 were selected for the study if clinicians across Genomic Medicine Centres (GMCs) reported a causative single nucleotide variant (SNV), insertion or deletion leading to a diagnosis. The AION pipeline was applied to 318 solved cases of which 57 were singletons and 261 had information from parents that could be used to guide the analysis. Subsequently, AION's performance was assessed using the variant prioritisation file produced for each case, ranked by statistical significance and the AION score.

AION's performance was assessed by calculating the percentage of cases where the diagnostic variant was identified. AION ranked the causative disease variant in the top 25 variants for 91.5% of cases. This means that an analyst can expect the disease-causing variant to appear in AION's prioritised output file for the majority of patients. When focusing on cases where parent information was available, AION prioritisation had a sensitivity of 93.1% (243/261). We observed even greater sensitivity of 94% (172/183) in the paediatric cohort and 96.8% (92/95) in the intellectual disability cohort, the most common rare disease category in the 100kGP. Importantly, the disease-causing variant had a median rank of 1 across all cases assessed. Therefore, in at least half of the cases analysed, the top-ranked variant by AION score was the diagnostic variant reported by clinicians across GMCs.

PATHOGENIC VARIANT IDENTIFIED - ETNICITY

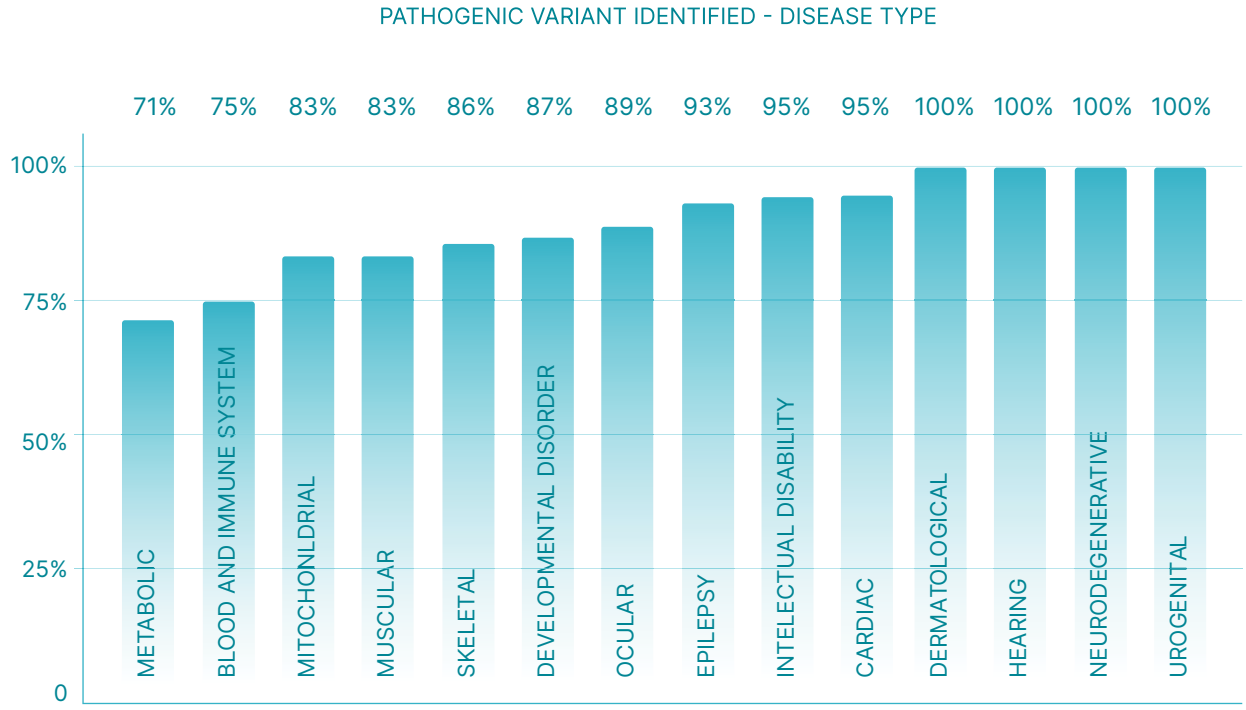


PATHOGENIC VARIANT IDENTIFIED - AGE GROUP



As the global genomic testing market expands with greater involvement of minority ethnic populations in genome sequencing research, it is critical to ensure that AI algorithms do not compound existing disparities in health outcomes. To determine AION's performance across diverse populations, we focused our analysis on patients with non-white self-reported ethnicity in the 100kGP. This filter revealed 39 patients eligible for analysis, among which AION ranked the pathogenic variant in the top 20 for 97% of cases. In comparison, performance was 89% for patients with white self-reported ethnicity and 93% for patients where ethnicity was not stated. While these results suggest that AION is not biased against diverse populations, this assessment is underpowered as ethnic minorities are under-represented in the 100kGP cohort. Future benchmarking studies will continue to monitor AION's performance across diverse populations.

Future work will expand on this phase 1 assessment of AION version 3.0 with a phase 2 study applying AION version 3.2 to 1,718 affected individuals (943 singletons, 775 trios) from the Pilot study of the 100kGP cohort. Benchmarking on our internal synthetic rare disease dataset shows significant performance increases in AION version 3.2, and our phase 2 study will compare these results to existing decision support tools used by Genomics England.



Conclusion

For rare disease patients to reap the benefits of genome sequencing, genomic testing services must overcome the variant interpretation bottleneck. Laboratories require technological innovations that improve diagnostic rates, streamline variant classifications, and support their limited personnel resources. At Nostos Genomics, we recognise that artificial intelligence has the power to address these challenges. AION is an AI-driven platform designed to prioritise genetic variants and streamline variant interpretation. A phase 1 study on the 100,000 Genomes Project Rare Disease Pilot cohort certifies that AION v3.0 identifies diagnostic variants with 91.5% sensitivity. Future benchmarks will see improvements in AION's sensitivity as scientists at Nostos Genomics continue to develop the powerful AI platform With AION, genomic services are liberated from the variant interpretation bottleneck and empowered on their journey to end the diagnostic odyssey for every patient.

To learn more about AION, book a demo or talk to our team.

References

[1] Nurk S, Koren S, Rhie A, Rautiainen M, Bzikadze AV, Mikheenko A, et al. The complete sequence of a human genome. Science. 2022;376: 44–53. doi:10.1126/science.abj6987

[2] Adams DR, Eng CM. Next-Generation Sequencing to Diagnose Suspected Genetic Disorders. N Engl J Med. 2018;379: 1353–1362. doi:10.1056/NEJMra1711801

[3] Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, 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: 405–424. doi:10.1038/gim.2015.30

[4] 100,000 Genomes Project Pilot Investigators, Smedley D, Smith KR, Martin A, Thomas EA, McDonagh EM, et al. 100,000 Genomes Pilot on Rare-Disease Diagnosis in Health Care - Preliminary Report. N Engl J Med. 2021;385: 1868–1880. doi:10.1056/NEJMoa2035790

[5] Liu P, Meng L, Normand EA, Xia F, Song X, Ghazi A, et al. Reanalysis of Clinical Exome Sequencing Data. N Engl J Med. 2019;380: 2478–2480. doi:10.1056/NEJMc1812033

[6] Eilbeck K, Quinlan A, Yandell M. Settling the score: variant prioritization and Mendelian disease. Nat Rev Genet. 2017;18: 599–612. doi:10.1038/nrg.2017.52

[7] Illumina Unveils Revolutionary NovaSeq X Series to Rapidly Accelerate Genomic Discoveries and Improve Human Health. In: Illumina [Internet]. 29 Sep 2022 [cited 23 Oct 2022]. Available: https://www.illumina.com/company/news-center/press-releases/press-release-details.html?newsid=8d04df3f-d9c1-4c85-8177-6ea604627ccd

Nostos Genomics

GET IN TOUCH

sales@nostos-genomics.com | www.nostos-genomics.com | book a demo

Nostos Genomics GmbH
c/o Stresemannstraße 123 Tenant GmbH,
Stresemannstraße 123,
10963 Berlin, Germany