

Artificial intelligence (AI) tools in genetics

 openaccessgovernment.org/article/artificial-intelligence-ai-tools-in-genetics/175533/

Vessela Kristensen and Dag Undlien uncover AI tools in genetics, from variant recognition to clinical implementation

Most people are curious about how their bodies work (and the ways they occasionally do not). This curiosity extends towards how our bodies are built, their functions, and what maintains life and health. Most people think that science is remote from the lives they lead, and the decisions that they make day by day, but this is far from the truth. Our understanding of genetics may affect our choices at our doctor's office about our healthcare and reproductive decisions, including family planning.

The connection between science and health

The connection between science and health is a direct one, and our ability to understand the science behind health affects our ability to understand our lives and our choices overall. For that, we need to study how specific mutations affect the cell, its compartments, the command centre, and the nucleus. Which functions are disrupted to cause disease and can be compensated with drugs or genetic repair? We employ studies in patients, advanced mathematical modelling and structural imaging to address these questions.

The complex human genome contains all the genetic information needed for development, functioning, and reproduction. Artificial intelligence (AI) and deep learning methods have been instrumental in extracting knowledge and patterns from this genomic data. Multiple genomic disciplines (for example, variant calling and annotation, disease variant prediction, epigenomics and pharmacogenomics, plus gene expression and regulation) take advantage of the generated high-throughput omics data (including next-generation sequencing, transcriptomics, proteomics, epigenetics), systematically annotated, curated, stored and organised in databases such as ENCODE, JASPAR or OmicsDI to utilise the power of deep learning algorithms to make predictions on the putative function of virtually every base in the genome out of the information stored there.

Why medical genetics?

The purpose of medical genetics is to daily attempt to identify and classify such genetic variants by the odds they are harmful, cause disease at different stages of life, and contribute to a particular drug response. The introduction of genomics has provided tremendous progress for patients with hereditary disorders.

However, dealing with the vast amounts of data generated makes it challenging to scale these kinds of analyses to ever-increasing demands in the healthcare system. A particular challenge is to sort out uncertainty in variant interpretation. So-called VUSs

(variants of uncertain significance) are abundant in the genome and a constant headache for variant interpreters, requiring lots of manual work.

Therefore, increasing the efficiency of clinical interpretation by cutting-edge solutions, including AI tools, is of pivotal importance for the scalability required to expand genomic medicine to all disease groups. The Nordic Alliance for Clinical Genomics brings together professionals across the Nordic countries to advance the clinical implementation of genomics. At its last annual meeting, the emerging AI-based tools were a central theme as they hold promise to boost variant interpretation. Recent and ongoing EU-regulatory activity that addresses some of the needs to facilitate this are: Informed consent and its harmonisation for clinical intervention, consent and dynamic consent (for data processing activities, including data sharing).

Genetic research ethics

Clinically, the EU regulation on a European Health Data Space, trustworthy AI and clinical practice address the challenges related to traceability and legal regulations. Genetic research ethics is a strong cultural and political focus through societal control and operational guidelines for ethical self-assessment of research in genetics and genomics. Insights into one's genetic background or interaction with machines can have both temporary and permanent effects, challenging the boundaries between health and illness, treatment and enhancement, normality and abnormality.

As geneticists, we are strongly committed to participating in this debate and contributing to the general acceptance of genetics and science as safe and responsible societal activities. The goal is to combine existing genetic knowledge with AI-based functional predictions. Computational and laboratory experiments are performed to create personalised risk predictions.

Machine learning-based prediction of variant function

Clinical genome diagnostics pipelines increasingly make use of whole-genome sequencing approaches. However, we cannot take advantage of most of these large data primarily because we lack an understanding of the functions, particularly non-coding regions. Based on how they influence the activity of the gene or gene networks implicated in the disease, we will predict new predisposing variants and test their ability to reclassify "variants of unknown significance" (VUS) by methods such as ExPecto, DeepSEA and SeqWeaver. Previously applied to autism, we believe they can be applied to cancer to reveal further how mutations in the non-coding genome can cause complex human diseases or disorders.

We can then confirm variants AI identified in our patient data by examining how they would have contributed to the initial diagnosis. This approach is particularly useful given that a large proportion of the diagnosed cases do not get a genetic explanation by

applying genetic testing with the current standards. It is also especially powerful for polygenic diseases such as cancer, psychiatric disorders, and autism, i.e., diseases with a strong genetic component.

Improving genetic diagnostic accuracy

This research will provide improved genetic diagnostic accuracy, facilitating more precise clinical follow-up of patients with ultra-rare diseases. Identifying genetic modifiers will be useful in providing personalised cancer risk prediction and optimising clinical follow-up for these patients and their family members. These findings will be important in identifying individuals who are more susceptible to developing cancer or disease at an earlier age and may aid in the development of strategies to prevent them in a personalised manner.



Please Note: This is a Commercial Profile



This work is licensed under [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/).