World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(REVIEW ARTICLE)

Check for updates

Exploring artificial intelligence for variant calling: A comprehensive review of tools and approaches

Mitali Dash * and Elamathi Natarajan

Biotecnika Info Labs, Bengaluru.

World Journal of Advanced Research and Reviews, 2026, 30(01), 1276-1290

Publication history: Received on 27 February 2026; revised on 06 April 2026; accepted on 09 April 2026

Article DOI: https://doi.org/10.30574/wjarr.2026.30.1.0877

Abstract

With the advent of Next-generation sequencing (NGS) methods, significant development is being observed in the identification of genetic variants. The precision and exactness of variant calls, nevertheless, continue to trouble researchers. Miscalculations can arise from multiple sources, such as sequencing artefacts, misalignments, and genomic heterogeneity. Conventional variant callers typically rely on statistical models that calculate the likelihood or probability of each variant call. However, they are not adept at handling messy data and mining variants from problematic genomic regions. To address such challenges, artificial intelligence (AI) is being explored. It is presented as the future of variant calling to enhance stability, precision and accuracy and eliminate errors.

The present review attempts to give a broad overview of AI-derived and AI-assisted variant calling approaches in trend. We have mentioned the drawbacks of traditional variant callers and described how artificial intelligence is being used. Following this, we have covered in length major AI-based tools such as DeepVariant, DeepTrio, Clair/Clair3, NeuSomatic, PEPPER–Margin–DeepVariant, Medaka, Hello and DNAScope.

We have provided an inclusive relative summary of the tools in use, highlighting the**ir** suitability, strengths, and drawbacks to aid in choosing the right tool. Additionally, we have suggested best practices to ensure the best output from AI-enabled pipelines. To give a complete, balanced outline, we detailed major constraints with AI-backed approaches, such as model transparency, overfitting and underfitting, computational requirements, and problems in structural variant detection. To sum up, we have offered a glimpse of emerging research areas such as transformer-based models, multi-modal data linkage, and GNN-based models.

Keywords: Artificial intelligence; Variant calling; Machine-learning; Deep learning; Neural Network; DeepVariant; Clair3; DeepTrio; NeuSomatic; PEPPER–Margin–DeepVariant; Medaka; Hello; Intelli-NGS

1 Introduction

Rapid advances in genome sequencing technologies have transformed disease diagnosis, precision medicine, and clinical therapeutics in the past decade. One of the critical reasons for the adaptation of next-generation sequencing is the ability to detect genomic variations. This, in technical terms, is called as **variant calling**, i.e. computationally identifying variations in a genome. These variations can be due to sequence alterations, such as single-nucleotide polymorphisms (SNPs), insertions and deletions (indels), s structural changes (large deletions, duplications, etc.) and copy number variations. The process of discovering genomic variants may appear relatively easy, but the execution part is complex and prone to errors from multiple sources, such as handling errors in the wet lab stage, instrumental errors during sequencing and software errors in the analysis stage. **(1)** Even though significant progress has been made in this direction, precise detection of variants remains a challenge.

* Corresponding author: Mitali Dash

The currently-popular variant-calling pipelines are largely based on parameterised statistical models and manual retuning. Frameworks such as Genome Analysis Toolkit (GATK) and FreeBayes have delivered stable results across wide-ranging scenarios. However, they falter in resolving tricky genomic regions, such as regions containing repetitive sequences and regions that have low sequence coverage. **(2,3).** Another point of concern with these conventional variant-callers is that they are well-adapted to short-read sequencing methods, which have a comparatively lower error rate. But the current trend shows a huge shift from short-read to long-read sequencing strategy. As such, long-read sequencing platforms tend to have higher error rates, which makes it challenging for the state-of-the-art variant callers to find variants efficiently. Hence, there is a need for new, advanced and adaptive tools to improve variant-calling results.

The integration of machine learning (ML) and deep learning (DL) into bioinformatics pipelines has enhanced the performance of variant-calling tools. The AI-driven tools can identify hidden patterns even in unprocessed data. This will help in the unearthing of real genetic variants, reducing false detections. While some AI-based callers, such as DNAScope, use simpler ML enhancements over conventional variant callers, the majority of AI tools rely on DL algorithms, specifically convolutional neural networks (CNNs). **(4)** Works are also being carried out using ML algorithms such as Stochastic Gradient Descent, Random Forest and XG Boost to build variant prediction models for cancer patients. **(5,6)**. AI-based variant callers such as DeepVariant, DeepTrio, and Clair3 are reported to show a higher precision rate over the existing variant callers across different sequencing platforms. **(7,8)**

Furthermore, these AI-enabled tools have been tested under various experimental conditions, yielding promising results. This suggests that AI tools can be flexibly efficient. There is also potential for scaling up variant-calling pipelines using AI. Yet, ML and DL focussed methods bring their own challenges. Some major obstacles limiting the adoption of AI-based tools are heavy computational demand, power and infrastructure requirements, difficulty in biological interpretation and the possibility of model bias. As AI-integration into actual clinical variant detection pipelines is not foolproof in the current context, there is a necessity to carry out meticulous comparisons for standardisation and performance validation. **(2,9,10)**

The current review attempts to cover a broad analysis of different AI-driven variant callers employed in the Next Generation Sequencing workflow. We have compared popular AI tools used in variant calling with respect to their strategy, performance, and cross-platform adaptability. We have presented our analysis of the evolving scenario in AI tools adaptation, as well as the possibilities and scope in AI-assisted variant detection.

2 Variant Calling Pipeline in Established Use:

Variant calling is vital for mining biologically and clinically relevant information from next-generation sequencing (NGS) data. It involves several sequential steps. Variant-calling tools analyse raw reads for quality issues, align them to a reference genome, perform post-alignment processing, and scan them to dig out genomic variants. It is important to proceed step-by-step to achieve high-quality calls. **(1)**

2.1 From Processing to Variant-Calling:

The initial quality control of raw sequences is intended to remove low-quality bases and trim the adapter sequences. **(1,11)** This ensures the inputs for downstream processes are reliable. Commonly used tools for quality control of sequence data include FastQC, MultiQC, Trimmomatic, CutAdapt, and Prinseq. Then, pre-processed reads are aligned to a reference genome, depending on the species choice. The algorithms commonly used for alignment include Burrows–Wheeler Aligner (BWA) or Bowtie2. **(1,12)** For speeding up the alignment process, lighter tools such as Salmon and Kallisto can also be used. The alignment step produces the sequence alignment map (SAM) files as output. The SAM file is later converted to the BAM file format. This makes the data computer-readable and improves storage efficiency. Proper alignment to the reference genome is key to success, as errors can negatively affect the results. **(1)**

2.2 Steps After Variant-Calling:

After the variant call, the post-processing step is equally important. This step mainly involves marking duplicate reads from PCR amplification, local realignment (specifically for indels), and base quality score recalibration (BQSR) for eliminating bias. **(1,13)**

2.3 Standard Variant-Callers

The state-of-the-art variant callers, such as GATK HaplotypeCaller, SAMtools/BCFtools, Strelka2, Varscan and FreeBayes, mainly employ statistical models. They follow rule-based methods that use parameters such as read depth, base quality, mapping quality, and allele frequency. **(2)** These features apply approximation rules to compute the

likelihood of variation at any specific location. The GATK HaplotypeCaller tool specifically targets active genomic regions where variation has been reported. It then reassembles the entire region into graphs, scans the reads and calculates the likelihood for specific haplotypes. **(1,13)** Evidence shows that a pipeline based on GATK Best Practices could finish a variant calling task on a 30X WGS sample in about 24 hours. **(14)**. In contrast, VarScan2 uses features such as read depth and allele frequency to identify new somatic mutations correctly. **(15)** Strelka2 also relies on a haplotype-based probabilistic model and is known to give accurate calls for detecting rare variants in cancer cases. **(16)** Both VarScan2 and Strelka2 are designed to be faster and more consistent than GATK. Nonetheless, Varscan2 is reported to be more reliant on feature adjustments. **(2,16)**

Similarly, SAMtools applies a Bayesian framework to compute the genotypic likelihood at every locus. It engages a multi-allelic model to detect genetic variants. **(17)** Compared to SAMtools, Free Bayes is more adaptive as it can handle data from polyploid species or mixed data from different individuals. But its results get affected by sequencing errors and feature thresholds. **(2, 15)**

2.4 Annotating the Variants to Know the Significance:

Annotation of the variants is the most crucial aspect, as this step helps to understand the biological relevance of the discovery and how they can be used to drive further research. **(3)** This step involves linking the variant-call results to different clinical and population databases, such as ClinVar and gnomAD. To get clearer and deeper insights into the biological impact of variants, functional annotation is performed.

2.5 Integrating AI into Existing Tools:

An important advancement seen these days is the integration of AI-based filtering techniques into traditional variant-calling pipelines. For instance, the Variant Quality Score Recalibration (VQSR) tool in GATK. This is an ML-based improvement that implements a Gaussian mixture model. **(1)** It groups genetic variants using different annotation criteria. Similarly, a random forest-based filter model is available in Octopus, another haplotype-based variant caller. GATK has also made available another improvement that implements a Convolutional Neural Network (CNN). This tool is known as CNNScoreVariants. **(3,6)** However, in all such cases, Machine-Learning or Deep-Learning has not been used for the actual variant calling. **(3)** While these algorithms enhance the accuracy and confidence, they are not completely independent and are linked to pre-designed conventions and assumptions. Hence, such methods are not very versatile nor complexity-compatible.

2.6 Challenges Associated with Conventional Variant Calling:

Though conventional variant callers are the tools of choice at present, they have not been able to achieve consistency and confidence. As they follow heuristic principles and offer only manual adjustments, the sharpness of detection is compromised in problematic portions of the genome. (2,4) Since they are largely trained and validated on short reads, their applicability to the longer reads becomes challenging. Another major concern of the conventional approaches is the difficulty in generalising the pipeline, as they are less flexible with newer sequencing technologies and experimental conditions. **(4)**

Furthermore, almost all the traditional variant-calling software is designed to work better in single-sample mode. Even though tools such as Octopus and Strelka2 sustain joint calling, only GATK is considered the ideal option for scaling up joint variant calling. **(3,15)** Besides, substantial perfections are required to fit the challenges of extensive WGS workflows. **(2,14)**

To address these problems observed in traditional variant-calling approaches, flexible, AI-driven tools have been designed to discover the hidden differences from unprocessed data. This will help in the better identification of variants for different applications. In the following section, we discuss well-known AI-based tools and their working principle.

3 Examining the Utility of Artificial Intelligence for Variant Calling:

The conventional variant-calling approaches are not ideal for complex genotypic analysis. They are prone to committing mistakes during the calling and pose obstacles to scalability and generalisation. **(2,3,4)** Their processing period is comparatively longer, which makes the entire pipeline inefficient and resource-intensive. **(2,14)** These constraints have forced scientists to embrace artificial intelligence (AI) to scale up, enhance precision, suitability, and robustness of the variant-calling process. **(4, 6)** Machine learning algorithms are typically incorporated to enhance the stability and sturdiness of the currently used variant-calling pipelines. With advances in technology, deep-learning architectures are now adapted to develop novel and hybrid tools. **(4,7,18)**

3.1 Application of Machine Learning

Machine learning algorithms are being exploited to devise feature-driven models for classifying genetic variants. The limits, however, are adjusted manually in most cases. Machine-learning classifiers, such as Support Vector Machines (SVMs), Random Forests, and Gaussian Mixture Models, utilise feature parameters like base quality scores, mapping quality, read depth, and strand bias to categorise variants. **(6)** Core Machine-learning models have not been used widely in the variant calling step specifically. Rather, they have been more commonly incorporated into the filtering and refinement stage. **(3,6)**

One such Machine-learning tool is the Variant Quality Score Recalibration (VQSR). This comes with the Genome Analysis Toolkit (GATK) and deploys a Gaussian Mixture model and computes confidence scores for individual variant calls. **(1,6)** Another commonly used tool is DNAscope, which incorporates a supervised machine learning algorithm into the GATK pipeline to increase the confidence in variant calls and enhance the pace of processing. **(19,20).**

Inclusion of machine learning has enhanced the consistency and sharpness of traditional methods. Yet, they need to match the rising level of complexity observed in emerging technologies and cloud architecture. **(3,6)**

3.2 Use of Deep Learning

Taking a cue from the success of earlier AI models, deep learning architectures have been tested for variant calling. Convolutional neural networks (CNNs) have become the most popular algorithm deployed in new and hybrid tools because they can learn underlying patterns easily. **(4,7)** Recently, Recurrent Neural Networks have become another option to choose from for variant calling. **(21,22)** Tools merging both CNNs and RNNs are also in vogue. **(8)**

DeepVariant, a widely appreciated tool since its release, utilises a CNN model to group the genomic variants **(7)**. It has been tested and validated on benchmarked sequencing reads for performance.

Another tool that improves upon CNN architecture is DeepTrio. This pulls out additional information from genetic patrimony sequences, making the variant calls sharper. **(23,24)** NeuSomatic is a specialised CNN-based tool targeted for cancer therapy. This can locate non-recurrent somatic variations easily. **(25)** Such approaches highlight the flexibility of deep learning models in integrating additional biological context.

Deep-learning models have also been customised for long-read sequences. Tools such as Clair and Clair3 have a hybrid architecture and are specifically designed for Oxford Nanopore and PacBio sequences. **(21,22,26)** Another recent tool in this domain is AlphaMissense from Google DeepMind, which employs a deep learning model to provide pathogenic interpretation of the variants. **(27)**

3.3 Methods of Conversion of Sequence Information using AI

The read data must be presented in the correct format to the AI algorithms for the best results. Depending on the working strategy, the encodings will vary. While some variant-callers are designed to work autonomously, many others depend on the associated software. There are two key methods to introduce the data to the algorithm:

3.3.1 Image Pileup Method

In this method, the reads after alignment onto the reference genome are first transformed into images and then interpreted by the algorithm as a two-dimensional matrix, called a tensor. In this matrix array, the horizontal organisation is meant for each unique read. The vertical arrangement refers to their genomic positions. **(4,7)** This method also considers other data, such as quality scores, strand directionality, and alignment information. This type of arrangement makes it easier for the model to interpret the data spatially, especially useful for the neural network-based tools**. (7)** These methods usually perform well on short-read sequencing data but falter with longer reads. Owing to the higher degree of miscalculations, supplementary cleaning steps are vital for such variable data.

3.3.2 Full Alignment Method

Another way to present data to the model is through full alignment. As the name indicates, this technique assembles and analyses alignment data, taking a broader frame of reference. The technique is better at scanning problematic areas, such as repetitive genomic regions, owing to greater awareness of the background data. **(8,26)** This technique, however, requires more computational power. **(8)**

Certain tools also combine the above methods to generate high-quality outputs. For instance, Clair3 uses a tensor-based pileup method for variant calling, followed by full alignment for adjustments and corrections. Similarly, PEPPER–Margin algorithm is fused with the DeepVariant algorithm to enhance confidence in outputs **(8,26).**

3.4 Merits of Applying AI Algorithms in Variant Calling:

Incorporating artificial intelligence into the variant calling process offers several benefits over conventional methods. The tools produce outputs that are relatively more accurate. They fare quite better in difficult-to-detect genomic regions. These models produce consistent outputs and are clutter-hardy. **(4,7,8)** Advanced tools can eliminate human-made faults by allowing the algorithms to select features. **(18)** The processing period and rate of variant calling have improved drastically with the inclusion of artificial intelligence. **(6,14)**

4 Established Variant Callers Leveraging Artificial Intelligence:

No single tool can be regarded as the best for variant calling when it comes to using advanced AI techniques. Different tools apply Artificial Intelligence in different ways, and hence their suitability varies according to the challenge handled. **(4,6)** Typically, variant callers are developed and customised to handle three broad categories of data – short reads, longer reads and somatic variations. In the following paragraphs, we provide a detailed overview of the working principles of well-known variant calling tools that employ machine learning and deep learning.

4.1 Variant Callers for Short Sequences

4.1.1 DeepVariant

DeepVariant is open-source software and has become the preferred tool over the past couple of years. This was originally released by Google Health and trained with a deep learning architecture. DeepVariant uses a convolutional neural network (CNN) to analyse aligned sequencing reads in the form of pileup images and find true variants. **(7)** DeepVariant often beats other available tools, such as GATK HaplotypeCaller, SAMTools, Strelka, and FreeBayes, in terms of precision and accuracy. **(7,8)** Its competence is mainly due to the automated filtration technique, which chalks out the need for additional fine-tuning. It was developed to work with short-read data, but it has been improved over time to work equally well for High-Fidelity (HiFi) longer reads. **(8)** However, one major downside of DeepVariant is its heavy computational needs. Even though it can be run on a Central Processing Unit (CPU), its speed improves significantly when connected to Graphics Processing Unit (GPU). **(6)**

4.1.2 DeepTrio

DeepTrio is an improvement over DeepVariant. Grounded in the same principle of neural networks, this tool integrates genome sequencing data from family trios, i.e. parent and offspring. It examines genomic data from three family members together. This tool exploits principles of Mendelian inheritance to reject faults, miscalculations, and directly determine variants or mutants afresh. As such, this helps to increase the veracity detection, especially for new and rare mutants. **(23,24)** Hence, DeepTrio is found to be appropriate for incorporating into clinical genomics pipelines. It has proven to be relatively superior to non-trio variant-callers such as GATK and FreeBayes **(23)**. Despite its potential, DeepTrio is restricted by the availability of heredity or trio data.

4.1.3 DNAscope

DNAscope is a tailor-made tool introduced by Sentieon to increase the processing speed and efficacy. **(19)** Initially launched targeting short-read sequences, DNAScope was later extended to handle both PacBio **(28)** and Oxford Nanopore sequences. **(29)**

DNAscope integrates Machine learning into the existing variant-calling pipeline of GATK HaplotypeCaller. This algorithm is trained on genotyping data and boosts the sensitivity and precision of the joint model. (19,20) DNAscope can handle large datasets faster and with precision, excluding manual refinement steps. It became quite successful in lowering operational costs by slashing the memory requirements and utilising multi-threaded processing. **(20)**

However, DNAscope cannot be regarded as a machine learning tool in the true sense. It is better considered as an ML-assisted AI tool. It does not require a GPU for scaling up, unlike other AI-derived tools.

4.2 Variant Callers for Longer Reads

4.2.1 Clair and Clairvoyante

Clair is a second-generation deep-learning variant calling tool. Clair is a successor to Clairvoyante. This is more versatile as it is customised to use both short and long sequences **(21)**. Clair implements a Recurrent neural network Bidirectional Long Short-Term Memory architecture (RNN-bi-LSTM) to analyse reads and yield truthful outputs. **(21,22)**. A major weakness of Clairvoyante was a higher error rate in multi-allelic variant calling. This was rectified in Clair by fine-tuning the model features and then training it on additional, assorted genomic data. **(21)** Even though Clair is validated to be superior for messy datasets, its efficiency is relatively lower for mining indels.

4.2.2 Clair3

Clair3 is the latest tool in the suite and entails a mixed architecture. It unites two methods of data analysis, i.e. pileup-based encoding and full-alignment-based processing. **(26)** This amalgamation helps to achieve greater reliability while reducing computational load. **(21)**

Clair3 is reported to have a relatively quicker run-time. It is known for producing consistent results in areas with lower sequencing coverage. **(26)** Clair3 has set a standard specifically for indel discovery.

4.2.3 PEPPER–Margin–DeepVariant

The PEPPER–Margin–DeepVariant is a unified tool intended for long-read sequences. This is a haplotype-aware variant caller and includes four major tools: **(4,8)**

1. PEPPER-SNP: This tool uses a recurrent neural network (RNN) to detect single-nucleotide polymorphisms (SNPs) from the aligned reads **(4,8)**

2. Margin: This tool is used at multiple stages in the pipeline for phasing and haplotyping. Firstly, Margin uses input data from PEPPER-SNP to produce an alignment file, i.e. haplotagged for the next step. It also combines DeepVariant results with the alignment data and organises them to produce a phased VCF file. Its working principle is based on the Hidden Markov Model (HMM). **(8)**

3. PEPPER-HP: It accepts the haplotagged outputs of Margin and scans them via a recurrent neural network. This predicts SNP and indel candidates. **(4,8)**

4. DeepVariant: DeepVariant analyses the predicted outputs of PEPPER HP. This data is integrated with the haplotagged alignment file from margin. Finally, DeepVariant executes the genotyping of likely variants, with reference to the given haplotype information.

The feature count is generally higher for DeepVariant than PEPPER, and its training is also comparatively efficient and accurate. But the disadvantage is the increased run time. Therefore, integration of PEPPER with DeepVariant can be a viable way to improve both speed and precision. **(4)**

This combinatorial approach is well-suited to analyse cluttered datasets, specifically to predict variants from structurally complex regions. This tool is unique as it is haplotype-aware. It can be used in pipelines where phasing is needed. But the tool is heavy on resource infrastructure.

4.2.4 Medaka

Medaka is developed by Oxford Nanopore Technologies, particularly for analysing sequence reads from ONT platforms. This tool uses aligned sequences and rectifies any mapping issues. The output is superior consensus sequences, which are supplied to the deep neural algorithm. **(30)** This deep-learning model is haplotype-aware and is trained on long-read ONT datasets. Hence, this tool is effective in mining small SNPs and indels while accounting for mistakes. Yet, its accuracy in finding bigger structural variants is not very high.

4.2.5 HELLO (Hybrid and stand-alone Estimation of small genOmic variants) ***(31)***

Hello is an open-source software customised for identifying SNPs and InDels from Illumina and PacBio separately and together. For Illumina sequences, it does a fresh alignment to improve the consistency of Indel mining. Similarly, for long reads, HELLO performs extra steps of haplotag sorting and alignment. HELLO uses a deep neural network

architecture to collate allelic evidence and then applies probabilistic reasoning for accurate variant calling. The network architecture of HELLO is relatively less intensive than that of DeepVariant. This permits the tool to run faster and limits computational load. **(6)**

4.2.6 Intelli-NGS (Intelligent NGS)

This tool was developed in a clinical setting, exclusively for handling the sequence data from Ion Torrent platforms. Intelli-NGS uses a deep neural network to rank genomic variants according to their associated confidence score.

The network architecture is trained to use VCF files. It sorts individual variants utilising the thirty-five features specific to the IonTorrent platform. Additionally, the source code allows annotation of every variant call through clinical databases such as dbSNP and ClinVar. Intelli-NGS was validated on GIAB datasets and reported to show an accuracy rate of 93.08% in variant delineation. The uniqueness of this model lies in its filtering approach, which considers the flow-space information, a parameter exclusive to Ion Torrent sequencers. This makes Intelli-NGS superior for Ion Torrent data in contrast to other variant callers, which ignore the flow-space parameter. **(32)**

*4.2.7 GARFIELD-NGS (Genomic vARiants FIltering by dEep Learning moDels in NGS) **(33)***

GARFIELD-NGS has the same purpose as Intelli-NGS, i.e. improving variant call quality by ranking them into high or low confidence groups. It is developed to work on Illumina as well as ION Torrent data. This tool was primarily tested on whole-exome sequencing, but it can also be applied to whole-genome sequencing. **(33)**

GARFIELD-NGS uses VCF files generated by GATK. It employs four deep learning models trained on INDELs and SNPs from Illumina and ION platforms. The model assigns probability scores called the CP value to variants and groups them accordingly. The tool works best on single-sample VCF files. It can take Multisample VCFs, but the confidence in predictions may not be as high as that of single-sample VCFs. **(33)**

4.3 AI Tools for Detecting Somatic Variants

4.3.1 NeuSomatic

NeuSomatic has specific applications in cancer genomics studies. This tool is known for identifying somatic mutations. (25) Neusomatic's model architecture includes convolutional neural networks and tensor encoding to transform aligned read data into arrays. This helps in distinguishing tumour sequences from normal ones. The model scrutinises specific tumour-normal read pairs and calls the variants.

NeuSomatic has been appreciated for its sharpness and perceptivity in detecting rare somatic variations. This tool is capable of handling heterogeneous cancer data. (25) Nevertheless, its dependency on paired datasets limits its application.

5 A Relative Evaluation of Performance of Popular AI-dependent Variant Callers

Though all AI-based variant callers have a similar objective, their working strategies and capabilities differ considerably. Their performance varies according to the sequencing platforms, types of variants targeted, and the clinical background. **(6)** Thus, it is essential to compare various variant-calling tools that are available. This will help in selecting the appropriate tool according to the objective.

5.1 Appraisal of AI-Backed Variant-Callers Concerning Different Sequencing Methods:

Each variant calling tool responds differently to different sequencing techniques, depending on its training data. Since short-read and hifi sequencing methods differ in the range of permissible faults, variant callers optimised for one method are bound to lag while handling sequences of other types. **(2,4)**

When assessing tools on popular short-read sequencers like Illumina, DeepVariant stands out. It outperforms other variant callers, such as Clair, due to its powerful deep neural network model that exploits tensor-based image pileup. **(7,8)** DeepTrio beats DeepVariant in terms of accuracy, as it incorporates patrimonial data to make more reliable variant calls. **(23,24)** The accuracy rate of DeepTrio in discovering SNPs can go up to 99.8% for short reads. It is even higher for PacBio sequences, touching 99.9%. **(7,8)** Hence, DeepTrio is singularly preferable for rare variant identification. **(23)**

In contrast, the strength of DeepVariant lies in its adaptability to multiple platforms. Its outcome is stable. The exactness of predicting SNPs and InDels remains quite high for both Illumina and PacBio sequences, often touching an F1 score of 99.9% **(7, 8)**. However, its performance drops while handling Oxford Nanopore sequence data. The fall is more noticeable in the case of InDels than in SNPs. **(8)** Thus, DeepVariant may not be the best choice for longer reads.

Clair3 is reported to be better than DeepVariant in handling longer reads, including Oxford Nanopore sequences. Since it unites the tensor representations method and the full alignment method, it can correct the sequencing mistakes. **(21,26)** It has been validated to show an accuracy rate of 99.9% for SNPs. Clair3 is also ideal for discovering InDels in genomic regions having lower sequence coverage. **(6)** Its faster run-time makes the tool a great option for research settings that deal with multiple sequencing methods. **(26)**

PEPPER–Margin–DeepVariant is specifically better at handling long read sequences. It often surpasses other tools such as Clair3 and Medaka for Oxford Nanopore data. As the tool is haplotype-aware, it can discover SNPs and Indels precisely, especially in regions containing homopolymers. **(8)**

Unlike many other tools, DNAscope is not open-source software. But it can be a better alternative to DeepVariant and DeepTrio as it demonstrates comparable efficiency for SNP discovery, and puts less pressure on computational infrastructure. DNAscope is tested to have an accuracy of about 99.9% for SNPs and 99.5% for InDels in PacBio long reads. Thus, DNAscope is useful in studies where processing time becomes important. **(19, 20)**

Similarly, Medaka is the ideal option for Oxford Nanopore reads. It is known for high-confidence SNP calls with an accuracy rate of about 99.0%. But it may not work very well for mining larger structural variants. **(6)** Besides, Clair3 outpaces Medaka when it comes to diploid variant calling owing to the model superiority. **(30)**

Hello's capability to succeed in research environments that tackle miscellaneous data makes it worth adopting. In hybrid setups such as Illumina-PacBio platforms, its detection accuracy for SNPs can be as high as 99.9%. Hello is also reported to considerably scale down the mistakes in InDel discovery. **(31)**. Hello fares better than DeepVariant in PacBio reads in both high and low frequency regions. **(31)**

5.2 Review of Tools Based on Their Capacity to Discover Somatic and Germline Variants:

DeepVariant and Clair3 are equally capable of discovering germline variants and perform better than most other variant callers. But their specificity varies according to the type of sequence reads. DeepVariant is undoubtedly the best for shorter reads. Clair3 is reliable for long-read datasets. **(7, 26)**

NeuSomatic is ideal for detecting somatic mutations. It is customised to delve deeper and find subtle non-conformities. This helps in segregating cancerous variants from normal ones. Compared with other variant callers, NeuSomatic excels at identifying rare somatic mutations associated with different cancers. **(25)**

5.3 Assessment of Variant Callers on the basis of Their Efficiency to Detect a Particular Type of Variant:

Most of the discussed tools do well in finding SNPs. However, their efficiency goes down while identifying InDels**. (2,4)** Only a few tools, such as DeepVariant and Clair3, have greater accuracy than conventional variant callers. **(7,26)** Merged frameworks, such as PEPPER–Margin–DeepVariant, are, however, levelling up over time to match these tools. **(8)**

5.4 Comparison of Variant Callers on the basis of Model Strength and Resource Use:

The major concern with AI-grounded tools is to balance model efficiency with resource use. DeepVariant has a very capable neural network model, but it needs huge computational resources. **(6,14)** In contrast, Clair3 has a lower power requirement with equal efficiency. Besides, Clair3 is comparatively less intensive on system infrastructure. Hence, it is more suitable in settings with limited availability of computational resources. **(26)** PEPPER–Margin–DeepVariant also has higher accuracy, but also higher system loads due to several integrative frameworks. **(8)** DNAscope can be a good option because of its ability to sustain the accuracy of variant-callers with a lesser computational load. **(20)**

5.5 Broad Comparison of Tools Using Common Parameters:

The above discussions suggest that there exists a lot of variability among the variant callers. It is difficult to compare them directly. None of the tools can be considered the best for every type of variant calling. It depends on the type of experiment, purpose, and infrastructural compatibility. **(2,4,6)**

To assist researchers in selecting the right tool, we designed the following table to summarise popular variant-callers.

Table 1 Comparison of Tools with Respect to Sequencing Platform, Use Case, Advantages and Limitations

Tool	Availability	Suitability	Sequence preference	Benefits	Constraints
DeepVariant	Opensource	Germline variant calling	Best for short reads, but can handle long reads	Higher rate of accuracy in calling SNP & indel, stable and reliable results	The model is heavy on the system. Slower processing time
DeepTrio	Opensource	Germline variant calling involving inheritance data	Better for short reads, but can handle long reads	The addition of Family history data makes the tool more relevant biologically. Good option for identifying rare variants.	Specialised need for parent-offspring or family trio data
Clair	Opensource	Early-stage variant caller	Both Long-read and short-read	System load is lower.	Less efficient than other tools
Clair3	Opensource	Long-read germline variant calling	Better for ONT/PacBio data, but supports shorter reads also	Good for scanning genomic regions with less coverage. Higher accuracy in finding indels on Oxford Nanopore reads. Faster run-time	It does not work well with very high coverage reads. It is not suitable for polyploid data
NeuSomatic	Opensource	Somatic variant calling, suitable for Cancer studies	both short and long read formats	The tool can capture variants that are less frequent.	Specialised need for normal-cancerous paired data
PEPPER–Margin–DeepVariant	Opensource	Scanning Complex genomic regions, phasing	Long-read	Haplotype-awareness; it can filter out clutter from real variants	Heavy system load
Medaka	Opensource	SNP mining in Oxford Nanopore sequence reads	Oxford Nanopore sequences only	Good at rectifying errors specific to Oxford Nanopore sequencing	Not good for finding InDels. Can not handle low-coverage data
DNAScope	Commercial	Good at detecting both SNPs and Indels, useful in cases where turnaround time is a constraint	Can handle both short read and long read data equally well	Faster run-time, computationally less intensive, gives reliable results in low coverage regions	It does not integrate the ML model fully. It is an ML-supported tool. It is not good at handling non-human sequences
Hello	Opensource	To handle assorted data coming from different platforms	Can work on both short-read and long-read data, specifically Illumina and PacBio	It adjusts and corrects the mistakes and artefacts in the supplied data	Not well validated and benchmarked for Oxford Nanopore sequences
Intelli-NGS	Opensource	Post Variant-calling filtration	Works for Short-reads	Improved accuracy and sharpness	Not very versatile Benchmarking is restricted

Garfield-NGS	Opensource	Variant-filtration	Compatible with Illumina and Ion-Torrent data	Enhances quality by removing false positives	Not a true variant caller

Table 2 Comparison of Tools with Respect to Supported Experiment Type, Data Encoding Technique, Model Architecture and Computational Compliance

Tool	**Type of Experiment allowed**	**Data Presentation Method**	**Network Model Employed**	**Compatibility with CPU/GPU**
DeepVariant	WGS, WES	Tensor-based image pileup	Convolutional Neural Network (CNN)	Both CPU (slow), GPU (ideal)
DeepTrio	WGS, WES (trio)	Tensor-based image pileup	Convolutional Neural Network (CNN)	CPU/GPU (GPU is better)
Clair	WGS	Tensor-based alignment summaries	Recurrent Neural Network (RNN) based on Bi-LSTM	Fully CPU compatible
Clair3	WGS (primarily), limited WES	A combination of array-based image pileup and full-alignment-based encryption	Two types of DL models: pileup-based neural network (CNN) + alignment model	Efficient on CPU, GPU non-compulsory
NeuSomatic	WGS, WES	Alignment feature matrices with different channels	Convolutional Neural Network (CNN)	GPU needed for best results
PEPPER–Margin–DeepVariant	WGS	Multi-level encryption: tensor encoding + alignment +image pile up	DL component includes RNN- Bi-directional GRU (Gated Recurrent Unit) + CNN	GPU is most critical
Hello	WGS, WES	Each read is treated as a sequence, not an image; pileup image encoding with the reference base; Full alignment encryption	Convolutional Neural Network (CNN) to work on input data and Dense layers for classification	Supported on CPU, GPU acceleration is helpful for better outcomes
Medaka	WGS	Tensor encoding with pileups obtained from alignment steps	Neural Network-based consensus model; mainly depends on RNN-LSTM	CPU-efficient, GPU not needed
DNAScope	WGS, WES	ML-assisted genotyping	Gradient-booster ML models	CPU-compliant GPU not needed

6 Best Practices to be Followed While Using AI Tools:

To get the best results from machine learning and deep learning-based tools, certain factors need to be considered cautiously. Though AI has advantages over conventional pipelines, the outcomes largely rely on the quality of the sequence data supplied to the model and the type of model chosen. From the above table, it is also clear that the operational cost of deploying a tool and the time requirements also need to be kept in mind. Hence, it is of prime importance to follow the best practices. This will guarantee that research output is high-quality and biologically relevant.

6.1 Ensuring Input Data Is of Good Quality and Bias-Free

The quality of the starting data is of utmost significance for the success of any variant calling pipeline. So, it is essential to comply with the standard practices, such as quality checking to eliminate low-quality bases, adapter traces, and other contaminants. **(1,11)**

Alignment to the specific reference genome should be done correctly to avoid misalignments that can introduce inaccuracies in the pipeline. **(1,12)** Processing of the aligned reads, such as sorting the BAM files, marking the PCR duplicates and indexing the reference genome, is also critical for AI-dependent pipelines. Otherwise, biases can persist in the data and negatively affect the predictions of AI models. (1)

6.2 Choosing the Right Variant Calling Tool as Per the Research Requirement and Availability

The second most important factor for the success of AI predictions is the selection of the right variant-calling tool. Here, the decision should be taken according to the research goal, the sequencing platform available, system infrastructure and computational power at disposal. For instance, DeepVariant may be the tool of choice if one works on the Illumina platform, wants to find germline variants and has access to good computational infrastructure. **(7)** On the other hand, for long-read sequencers with the same agenda as before, there are two options: Clair3 or PEPPER–Margin–DeepVariant. **(26, 8)** If computational capacity is a constraint, then Clair3 is the best choice among the two. **(26)**

Similarly, if a researcher wants to incorporate pedigree information into the pipeline to accurately discover novel and rare mutations, then DeepTrio should be the preferred tool. **(23)**

These facts highlight that choosing the right variant caller is not straightforward. Rather, it should accommodate the research objective, all resources, and constraints to avoid any delays or errors in pipeline execution. **(6)**

6.3 Sequence Coverage:

The machine-learning or deep-learning models are generally trained on datasets within a particular coverage span. Hence, they may not do well beyond those limits. The models may fail to identify variants, or there may be errors in calculation. **(6, 34)** Thus, uniform coverage is vital to increase the precision of model predictions.

6.4 Need for Computational Power and Infrastructural Facilities:

Since AI-dependent tools, specifically those that utilise deep neural networks, can be heavy on the system infrastructure and computational resources, it is essential to select tools based on feasibility. **(6,14)** Sometimes, GPU acceleration is required for the best result, for example, DeepVariant. The key factor for success is to achieve the best possible trade-off between output quality and cost. If there are any constraints, alternatives should be adopted. **(6,20)** For example, Clair3 can be a good option than DeepVariant if computational power is restricted.

6.5 Validation of Results:

To find out potential issues with variant calls done using any tool, the results should be tallied with different datasets. Independent datasets can also be exploited to validate the results. **(35)**

6.6 Reproducibility and Workflow Standardisation:

Maintaining reproducibility is a major challenge in variant calling pipelines that use AI. **(3)** To overcome this, recording the versions, configurations, and dependencies of tools used is crucial. The features used in the pipeline should be clearly mentioned. **(1)** Container platforms such as Docker can be used to maintain the stability of the developed pipeline on different systems. Automation packages such as Nextflow or Snakemake can also be incorporated to make the pipeline transferable and replicable.

6.7 Linking Variant callers with Downstream Applications:

The variant calling pipelines need to be integrated with downstream steps, such as the annotation process, to assess the real biological relevance of output data. **(36,5)** Therefore, to further enhance interpretability, variant callers should be linked to online clinical databases and servers.

7 Constraints of Variant Callers that Work on Principles of AI:

Though AI techniques have progressed much, AI-derived variant-calling tools are confronted by many technical and operational obstacles that hinder their direct adoption in research and clinical pipelines.

7.1 Non-transparent Mechanism of Decision-Making:

As the tools discussed predominantly use AI for prediction, they suffer from the typical **black-box syndrome**. Conventional variant-calling tools allow users to understand how predictions are made through statistical measures. In contrast, neural networks usually hide the internal decision-making from the users. This makes the model classification questionable, in turn leading to difficulties in regulatory sanctions **(6, 7, 9, 10, 37)**.

7.2 High Influence of Training Data and Obstacles in Model Generalisation:

The predictions made by AI models are determined by the nature of their training data sets. If the data is of good quality, then the results are expected to have higher confidence. **(4,35,38)** The diversity of training datasets also matters, as tools otherwise can miss out rare variants and novel mutations in the actual data. AI models are also vulnerable to biases existing in training data. This can cause either underfitting or overfitting, leading to inaccuracies in the downstream process. Consequently, models trained on less-than-ideal datasets are difficult to generalise. **(4,39)** Hence, results may vary according to the source of genomic material, sequencing reads and the type of experiment.

7.3 Inability to Identify Structural Variants:

Most of the variant callers discussed above can mine SNPs and indels with greater accuracy and consistency. But they struggle when it comes to finding **Structural Variants**. Since SNPs and Indels are smaller and the extent of variation is less, models can predict with precision. But, more complex variants, such as large insertions and deletions, are hard to forecast correctly. **(4, 8)**

7.4 Computational Power Demand:

In general, AI-dependent variant callers need greater computational power. Some of the tools perform better if given GPU support. For example, DeepVariant and NeuSomatic. This ultimately leads to increased energy consumption and pressure on infrastructure. Therefore, in resource-limited settings, AI adoption can be challenging. **(6,7,8)**

7.5 Difficulty in Adoption by Clinical Setup:

Direct adoption of AI tools in clinical pipelines is not yet preferred. The reasons include problems in proper validation, benchmarking, and monitoring. **(9)** Clinical research environments need tools that can ensure accuracy and precision while making the pipeline replicable. AI models in use cannot satisfactorily fulfil these demands. It is difficult to integrate AI into existing pipelines because of the requirement for complete refurbishment. **(2,3)**

Considerable improvement is observed in variant-calling outcomes due to the inclusion of AI technology. But they come with their own set of challenges. These challenges are to be overcome to facilitate their integration into real-world projects.

8 Future Scope for AI-Enabled Research:

AI-derived technologies are continuously changing, and researchers aim to enhance the accuracy and precision of model predictions, interoperability, and scalability. Simultaneously, new technological trends are also emerging. So, they can transform the future of variant calling to make it more versatile and comprehensive.

8.1 AI Models with Transparent Internal Mechanisms:

The mechanism adopted by AI needs to be explainable to ensure transparency within the pipeline. Hence, a principal area of research is to enhance the observability of the working mechanisms of neural network models. Focusing on improving the attention process and explainable AI (XAI) frameworks can help to ensure trust in AI-predicted variant-calling results. This will benefit clinical researchers who want to use AI-based tools**. (9,10,38)**

8.2 Linking the Model Predictions to Multi-Modal Data:

Next-generation variant callers will have to handle more complex datasets from multiple sources. Hence, they need to be more advanced to track underlying inferences from various types of data, such as transcriptomics, epigenomics,

inheritance data, and linking these modalities to combined workflows. This will increase biological relevance and understanding of variant calls. **(18)**

8.3 Exploring Transformer-driven Network Models:

The evolution of **transformers** has paved the way for developing new frameworks that can handle challenges inherent to CNN or RNN-based approaches. Transformers can be extremely useful for uncovering patterns from large and complex genomic regions, increasing the chances of discovering difficult-to-capture variants. **(18, 40)**

8.4 Shift to Graph-Neural Network-Based Models:

Graph Neural Network (GNN) models are noted for their benefits over Convolutional Neural Networks in handling complex and networked data. Since GNNs have the capacity to interpret graph-based reference genomes, they can identify any variations more easily than linear models**.** GNNs provide benefits over CNNs in resolving sparse datasets. Hence, research on GNN-backed tools has the potential to take variant calling to a new level. **(41)**

8.5 Switching to Combined Transformer Models for Structural Variant Detection:

An emerging research trend is to integrate CNN-based models with transformers for boosting the structural variants discovery rate. Extending research to tackle long-read sequencing data with greater complexity is the key goal. **(4,18)**

8.6 Adapting Cloud-Based Architecture:

To resolve the issues arising due to resource and infrastructural limitations, research is being directed towards developing variant calling tools that can be run entirely on cloud networks. Tools that benefit from **distributed computing infrastructures** also need to be explored. This will help in scaling up the variant calling pipelines and promote data sharing for research collaboration. **(42)**

8.7 Summary:

The next wave of advancements in AI-associated technologies is going to alter the current method of variant calling significantly. The objective is to make the tools more transparent, consistent, replicable, and generalised. New research can considerably enhance the accuracy and quality of variant calling tools, leading to wider adoption in genomics research.

9 Conclusion

Inclusion of Artificial Intelligence has led to greater efficiency and consistency in variant calling outcomes. Identification of indels, variants from tricky areas, and rare mutants has seen a quantum jump owing to tools such as DeepVariant, Clair3, NeuSomatic, and PEPPER–Margin–DeepVariant. Yet, challenges persist due to the inherent drawbacks of AI models. None of the models discussed can be singularly considered as the most ideal for all use cases. The prediction outcomes differ according to the sequencing platform, data quality, and research objective. Hence, tools should not be selected randomly; rather, tools should be chosen after careful consideration of all operational and experimental conditions. It is also recommended to follow the established best practices to increase confidence in research results. AI tools are still not trusted to be adopted in sensitive clinical research settings. Future research centred on Graph-Neural-Networks, Transformers and Cloud computing will augment the capacity and practical utility of AI-backed variant-calling tools. Emerging research has the potential to eliminate limitations and promote the adoption of AI-enabled frameworks.

Compliance with ethical standards

Disclosure of conflict of interest

The authors declare that they are not bound by any financial or non-financial agreements or relationships that will create a potential conflict of interest.

Use of Generative AI

The authors declare that Generative was used to assist in drafting and editing this review. However, scientific arguments and analyses were made and verified by the authors to uphold the novelty of the article.

References

[1] Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, et al. From FastQ data to high-confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics.* 2013;43(1):11.10.1–11.10.33. doi:10.1002/0471250953.bi1110s43

[2] Zhao S, Agafonov O, Azab A, Stokowy T, Hovig E. Accuracy and efficiency of germline variant calling pipelines for human genome data. *Sci Rep.* 2020; 10:20222. doi:10.1038/s41598-020-77218-4

[3] Barbitoff YA, Ushakov MO, Lazareva TE, Nasykhova YA, Glotov AS, Predeus AV. Bioinformatics of germline variant discovery for rare disease diagnostics: current approaches and remaining challenges. *Brief Bioinform.* 2024;25(2): bbad508. doi:10.1093/bib/bbad508

[4] Ren J, Zhang Z, Wu Y, Wang J, Liu Y. A comprehensive review of deep learning-based variant calling methods. *Brief Funct Genomics.* 2024; 23:303–313. doi:10.1093/bfgp/elae003

[5] Doig KD, Perera R, Kankanige Y, Fellowes A, Li J, Lupat R, et al. Using artificial intelligence to model clinical variant reporting for next-generation sequencing oncology assays. *BioData Min.* 2025;18(1):74. doi:10.1186/s13040-025-00489-y

[6] Abdelwahab O, Torkamaneh D. Artificial intelligence in variant calling: a review. *Front Bioinform.* 2025; 5:1574359. doi:10.3389/fbinf.2025.1574359

[7] Poplin R, Chang PC, Alexander D, Schwartz S, Colthurst T, Ku A, et al. A universal SNP and small-indel variant caller using deep neural networks. *Nat Methods.* 2018;15(10):821–828. doi:10.1038/s41592-018-0167-z

[8] Shafin K, Pesout T, Chang PC, Nattestad M, Kolesnikov A, Goel S, et al. Haplotype-aware variant calling with PEPPER–Margin–DeepVariant enables high accuracy in nanopore long reads. *Nat Methods.* 2021;18(11):1322–1332. doi:10.1038/s41592-021-01299-w

[9] Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med.* 2019;25(1):44–56. doi:10.1038/s41591-018-0300-7

[10] Samek W, Wiegand T, Müller KR. Explainable artificial intelligence: understanding, visualizing and interpreting deep learning models. *Proc IEEE.* 2017;105(10):1971–1990. doi:10.1109/JPROC.2017.2676999

[11] Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible read trimming tool for Illumina sequence data. *Bioinformatics.* 2014;30(15):2114–2120. doi:10.1093/bioinformatics/btu170

[12] Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012;9(4):357–359. doi:10.1038/nmeth.1923

[13] McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analysing next-generation DNA sequencing data. *Genome Res.* 2010; 20:1297–1303. doi:10.1101/gr.107524.110

[14] Liu H, Zhang Q, Ai F, Bu F. Benchmarking germline variant calling performance of a GPU-accelerated tool on whole-genome sequencing datasets. *Research Square* [Preprint]. 2024. doi:10.21203/rs.3.rs-4318731/v1

[15] Kim S, Scheffler K, Halpern AL, Bekritsky MA, Noh E, Källberg M, et al. Strelka2: fast and accurate calling of germline and somatic variants. *Nat Methods.* 2018;15(8):591–594. doi:10.1038/s41592-018-0051-x

[16] Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* 2012;22(3):568–576. doi:10.1101/gr.129684.111

[17] Garrison E, Marth G. Haplotype-based variant detection from short-read sequencing. *arXiv* [Preprint]. 2012

[18] Eraslan G, Avsec Ž, Gagneur J, Theis FJ. Deep learning: new computational modelling techniques for genomics. *Nat Rev Genet.* 2019;20(7):389–403. doi:10.1038/s41576-019-0122-6

[19] Freed D, Pan R, Chen H, Li Z, Hu J, Aldana R. DNAscope: high accuracy small variant calling using machine learning. *bioRxiv* [Preprint]. 2022. doi:10.1101/2022.05.20.492556

[20] Glusman G, Margarido GR, Aganezov S, Mainzer LS, Kendig KI, Baheti S, et al. Sentieon DNASeq variant calling workflow demonstrates strong computational performance and accuracy. *Front Genet.* 2019; 10:736. doi:10.3389/fgene.2019.00736

[21] Luo R, Sedlazeck FJ, Lam TW, Schatz MC. Clair: exploring the limit of using a deep neural network on pileup data for variant calling. *Bioinformatics.* 2019;35(20):4168–4175. doi:10.1093/bioinformatics/btz207

[22] Luo R, Sedlazeck FJ, Lam TW, Schatz MC. A multi-task convolutional deep neural network for variant calling in single molecule sequencing. *Nat Commun.* 2019; 10:998. doi:10.1038/s41467-019-09025-z

[23] Kolesnikov A, Goel S, Nattestad M, Yun T, Baid G, Yang H, et al. DeepTrio: variant calling in families using deep learning. *Nat Commun.* 2021; 12:563. doi:10.1038/s41467-021-20972-6

[24] Brand F, Guski J, Krawitz P. Extending DeepTrio for sensitive detection of complex de novo mutation patterns. *NAR Genom Bioinform.* 2024;6: lqae013. doi:10.1093/nargab/lqae013

[25] Sahraeian SME, Liu R, Lau B, Podesta K, Mohiyuddin M, Lam HYK. Deep convolutional neural networks for accurate somatic mutation detection. *Nat Methods.* 2019;16(10):991–999. doi:10.1038/s41592-019-0494-4

[26] Zheng Z, Li S, Su J, Leung AW, Lam TW, Luo R. Symphonizing pileup and full-alignment for deep learning-based long-read variant calling. *Nat Methods.* 2022;19(7):743–751. doi:10.1038/s41592-022-01543-0

[27] Cheng J, Novati G, Pan J, Bycroft C, Žemgulytė A, Applebaum T, et al. Accurate proteome-wide missense variant effect prediction with AlphaMissense. *Science.* 2023;381(6664): eadg7492. doi:10.1126/science. adg7492

[28] Olson ND, Wagner J, McDaniel J, Stephens SH, Westreich ST, Prasanna AG, et al. PrecisionFDA Truth Challenge V2. *Cell Genomics.* 2022; 2:100129. doi: 10.1016/j.xgen.2022.100129

[29] Sentieon. DNAscope LongRead nanopore pipeline [Internet]. 2025. Available from: https://www.sentieon.com

[30] Oxford Nanopore Technologies. Medaka: sequence correction tool [Internet]. 2018. Available from: https://github.com/nanoporetech/medaka

[31] Ramachandran A, Lumetta SS, Klee EW, Chen D. HELLO: improved neural network architectures for small variant calling. *BMC Bioinformatics.* 2021; 22:404. doi:10.1186/s12859-021-04311-4

[32] Singh A, Bhatia P. Intelli-NGS: deep neural network-based variant filtering. *Research Square* [Preprint]. 2020. doi:10.21203/rs.3.rs-1539740/v1

[33] Ravasio V, Ritelli M, Legati A, Giacopuzzi E, Della Mina E, Barlati S, et al. GARFIELD-NGS. *NAR Genom Bioinform.* 2020;2(1): lqaa007. doi:10.1093/nargab/lqaa007

[34] Sims D, Sudbery I, Ilott NE, Heger A, Ponting CP. Sequencing depth and coverage. *Nat Rev Genet.* 2014;15(2):121–132. doi:10.1038/nrg3642

[35] Zook JM, Chapman B, Wang J, Mittelman D, Hofmann O, Hide W, et al. Benchmark SNP and indel genotype calls. *Nat Biotechnol.* 2014;32(3):246–251. doi:10.1038/nbt.2835

[36] Calvino G, Farro J, Zampatti S, Peconi C, Megalizzi D, Trastulli G, et al. From genomics to AI. *Appl Sci.* 2025; 15:6578. doi:10.3390/app15126578

[37] Poushter J, Fagan M, Corichi M. How people around the world view AI [Internet]. Pew Research Center; 2025. Available from: https://www.pewresearch.org

[38] Wilk EJ, Taluri S, Howton TC, Crumley AB, Mrug M, Lasseigne BN. AI in variant analysis. *Preprints.* 2026. doi:10.20944/preprints202603. 1025.v1

[39] Lazer D, Kennedy R, King G, Vespignani A. The parable of Google Flu. *Science.* 2014;343(6176):1203–1205

[40] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. *Adv Neural Inf Process Syst.* 2017; 30:5998–6008

[41] Garrison E, Sirén J, Novak AM, Hickey G, Eizenga JM, Dawson ET, et al. Variation graph toolkit. *Nat Biotechnol.* 2018;36(9):875–879. doi:10.1038/nbt.4227

[42] Poplin R, Chang PC, Alexander D, Schwartz S, Colthurst T, Ku A, et al. Scaling variant discovery. *bioRxiv* [Preprint]. 2018. doi:10.1101/201178