AlphaMissense: AI predictions of genetic variant pathogenicity in clinical genetics

IMPLEMENTED 2023–2025 · HEALTH AND LIFE SCIENCES · GLOBAL · BAND D · PARTLY PROVED
VERSION 1.0 · PUBLISHED 20 NOVEMBER 2026 · UPDATED 20 NOVEMBER 2026 · FACTS CHECKED 20 NOVEMBER 2026

Google DeepMind released AlphaMissense in 2023 with predictions for all 71 million possible missense variants, and EMBL-EBI built the scores into Ensembl, UniProt, DECIPHER, ProtVar and the AlphaFold Database in 2024. The ClinGen working group calibrated it for clinical variant classification in 2025: it can add evidence at up to Strong level, but only at stricter thresholds than the developer proposed and with at best modest gains over four tools already recommended. UK guidance in 2024 noted that REVEL is commonly used in UK labs, and a 2025 rare disease study found 32.9% precision against expert-curated pathogenic variants.

ABOUT THIS SYNOPSIS

This is MarLuc's summary and grading of publicly available sources, for information only. It is not investment, legal, financial or professional advice, and should not be relied on for any decision. Readers must validate it against the original sources and, where appropriate, challenge its conclusions.

Facts and figures are as reported by the sources listed below, which may be incomplete, self-reported or out of date. Check the original source before relying on any figure.

The impact band, proof marker and SEAM view are MarLuc's opinion. The organisation is not a MarLuc client and has not reviewed or endorsed this synopsis. Names and trademarks belong to their owners.

NOTICE WORDING LAST UPDATED 16 NOVEMBER 2026

AT A GLANCE

Frontier case: AI used for a scientific, medical or engineering breakthrough. MarLuc grades what has changed in practice so far.

AI used: Deep-learning model adapted from AlphaFold and fine-tuned on human and primate population variant data; gives each possible single amino acid change a pathogenicity score between 0 and 1.

Implemented: 2023–2025. Grade: band D, Task support, Partly proved.

Headline result, as reported: AlphaMissense scores all 71 million possible missense variants and is built into EMBL-EBI tools; ClinGen's 2025 calibration found it can give up to Strong evidence of pathogenicity at stricter thresholds, with at best modest gains over existing tools, and the developer's own threshold did not reach even Supporting evidence

1 · The case

The facts, as reported by the 5 sources listed at the end. Where a source is silent, the field says "Not reported".

Organisation

Google DeepMind, with EMBL-EBI; calibrated for clinical use by the ClinGen Sequence Variant Interpretation Working Group. Google DeepMind, Alphabet's AI research lab, released AlphaMissense as a public resource. Clinical use is governed by laboratory standards: the ACMG/AMP variant classification rules, ClinGen recommendations in the US and ACGS guidelines in the UK. Canvas block: 6. Data and technology. Also 4. Operations and process.

Business challenge

The challenge. Per DeepMind, of more than 4 million missense variants seen in humans, only 2% have been annotated as pathogenic or benign by experts; the rest are variants of unknown significance. Per ClinGen, the original ACMG/AMP rules limited evidence from computational tools to the weakest strength.

How it was addressed. DeepMind fine-tuned AlphaFold on variants seen in human and primate populations, treating common variants as benign, and released scores for all possible variants with open model code. EMBL-EBI built the scores into its public resources. The ClinGen Sequence Variant Interpretation Working Group then calibrated the scores against ClinVar pathogenic and benign variants to set thresholds for each ACMG/AMP evidence strength. [1] [2] [3] [4] [5]

Who implemented it

Google DeepMind (Model or cloud provider): Developed AlphaMissense, published the predictions and open-sourced the model code..
EMBL-EBI (University or research lab): Integrated the scores into Ensembl, UniProt, ProtVar, DECIPHER and the AlphaFold Database..
ClinGen Sequence Variant Interpretation Working Group (University or research lab): Independently calibrated the scores for clinical PP3/BP4 evidence..
Genomics England (Industry partner): Cross-referenced the predictions with its variant pathogenicity data, per DeepMind..
Clinical laboratories that have adopted AlphaMissense as their pre-specified tool are not named in the sources opened.

Technology (SEAM pillar E)

SEAM types: T5 Custom / fine-tuned models, T8 Predictive / ML models. Deployment: Not reported. Vendors and products: Google DeepMind, AlphaMissense, Ensembl Variant Effect Predictor, UniProt, DECIPHER, ProtVar, AlphaFold Database.

Where it was built in: Distributed as precomputed scores, open model code and integrations in EMBL-EBI resources, including the Ensembl Variant Effect Predictor web interface, REST API and command line, which laboratories use to annotate variants before classification.

What changed

Before: ClinGen had calibrated four tools (BayesDel, MutPred2, REVEL and VEST4) for computational evidence up to Strong. After: AlphaMissense sits inside standard variant annotation tools, and ClinGen thresholds let it provide evidence up to Strong for pathogenicity and Moderate for benignity. The classification process is unchanged: computational evidence remains one input among several. [1] [2] [3] [4] [5]

What it cost

Not reported.

What it saved

Not reported. No time or cost saving for laboratories, and no change in diagnostic yield, is reported in the sources opened. [1] [2] [3] [4] [5]

Governance and control

DeepMind states the predictions are not for direct clinical use. ClinGen calibrated them independently and recommends one pre-specified tool per laboratory or gene; UK guidance recommends PP3 at Supporting strength only, pending ACMG updates.

S2 · Accountable owner · Claimed
Google DeepMind developed and released the model and predictions; EMBL-EBI runs the resources that integrate them. [1] [2]

A1 · Safe use · Claimed
DeepMind says the predictions are not designed to be used in the clinic directly and should be interpreted with other sources of evidence. [1]

A2 · Human in the loop · Evidenced
Clinical classification under ACMG/AMP rules always requires several lines of evidence, and UK guidance recommends PP3 at Supporting strength only, so the score cannot decide a classification alone. [3] [4]

A3 · Monitoring and assurance · Evidenced
ClinGen calibrated the scores independently against ClinVar, and an independent cohort study measured precision and recall against expert-curated classifications. [4] [5]

A4 · Vendor AI · Evidenced
The model code is open source and the predictions are under a CC BY 4.0 licence; ClinGen states new tools must be independently calibrated before clinical use. [1] [4]

A5 · Agent controls · Not reported
No agents described.

A6 · Standards and regulation · Evidenced
ClinGen's calibration maps scores to ACMG/AMP evidence strengths; the ACGS 2024 UK guidelines, ratified on 20 February 2024, set how in silico evidence is used in UK laboratories. [3] [4]

What worked, what did not, what is unproven

WORKED

– At calibrated thresholds AlphaMissense reached the Strong level of evidence for pathogenicity and Moderate for benignity, similar to the four tools ClinGen already recommended.

– Its scores are built into Ensembl, UniProt, ProtVar, DECIPHER and the AlphaFold Database; EMBL-EBI says the integration responded to requests from its user communities.

– DeepMind lifted the non-commercial restriction on the predictions in March 2024.

– In ClinGen's analysis, AlphaMissense gave higher likelihood ratios in the Strong pathogenic interval than the four earlier tools, though the authors say this may reflect the small number of

DID NOT

– The developer's recommended threshold (0.564) did not meet even the Supporting level of evidence in ClinGen's calibration; stricter thresholds were needed.

– In a Boston Children's Hospital rare disease cohort, precision was 32.9% and recall 57.6% for expert-curated pathogenic variants, and the model struggled in intrinsically disordered protein regions.

– ClinGen found at best modest improvements over the previously recommended tools, and said AlphaMissense's use of the term likely pathogenic is inconsistent with clinical classification terminology.

UNPROVEN

– Whether using AlphaMissense increases diagnoses or changes patient care: no outcome data in the sources opened.

– How many clinical laboratories use it as their pre-specified tool: Not reported.

– Performance gene by gene; ClinGen lists limitations including potential circularities and gene- or disease-specific variability.

– Whether training on gnomAD variants as a proxy for benign variants affects results; the ClinGen authors say it is unclear.

Timeline

Sept 2023 Google DeepMind publishes AlphaMissense in Science and releases predictions for all 71 million possible missense variants [1]

Feb 2024 ACGS ratifies its 2024 UK variant classification guidelines, which describe AlphaMissense as a new tool and note REVEL is commonly used in the UK [3]

Mar 2024 DeepMind moves the predictions to a CC BY 4.0 licence, lifting the non-commercial restriction [1]

May 2024 EMBL-EBI integrates AlphaMissense into Ensembl, UniProt, ProtVar and the AlphaFold Database, adding DECIPHER in June [2]

Feb 2025 npj Genomic Medicine study reports discordance with clinical-grade classification in a rare disease cohort [5]

Mar 2025 ClinGen working group publishes calibrated AlphaMissense thresholds for PP3/BP4 evidence in Genetics in Medicine [4]

2 · Grade and SEAM assessment

MarLuc's opinion, made on the evidence above. How cases are graded: marluc.co.uk/case-studies/how-we-rate

Level of change and impact

Claimed: Not applicable. Grade: D, Task support. Proof: Partly proved. Level of change: Task support.

Task support (Level 1, band D): variant scientists gain one more calibrated evidence source within an unchanged classification process. Not graded higher because no source shows

The SEAM view and lessons learnt

Availability came from the developer's open release and EMBL-EBI's integration; clinical acceptance depended on independent calibration, which found useful but modest gains and thresholds stricter than the developer's. The AI's effect on diagnoses cannot be separated, because the rules require several lines of evidence and no outcome study is reported. Before adopting an AI score in a regulated decision, a leadership team should ask who calibrated it independently against its own evidence standard, and whether it measurably improves results over the tool already in use.

1. Calibrate any AI score independently before using it in a regulated decision.
2. Check the developer's thresholds against your own evidence standard before adopting them.
3. Measure whether a new tool improves results over the one it would replace.

Evidence

Score 2 of 4. Independent: peer-reviewed calibration by the ClinGen working group (Genetics in Medicine 2025) and an independent cohort study (npj Genomic Medicine 2025); UK professional guidance (ACGS 2024). Developer results rest on its own Science paper. No outcome study.

Evaluation method. Independent evaluation: retrospective calibration against ClinVar pathogenic and benign variants (ClinGen) and comparison with expert-curated classifications in a rare disease cohort (Boston Children's Hospital). No prospective or outcome study.

Sources disagree. DeepMind reports classifying 89% of variants at a threshold with 90% precision on known disease variants; ClinGen found the developer's threshold did not reach Supporting evidence, and the Boston cohort study found 32.9% precision. The studies use different data, and the cohort authors note that pathogenic variants are rarer in real-world data than in balanced benchmark sets.

Status. In use as a research and evidence resource; calibrated thresholds for clinical PP3/BP4 use published in March 2025. The 2024 UK guidelines describe it as a new tool without recommending it as the default.

3 · Original sources

Read these in full. To challenge a fact or a grade, use the contact form at marluc.co.uk or the link on https://marluc.co.uk/case-studies/google-deepmind-alphamissense-variant-classification

[1] Google DeepMind: AlphaMissense catalogue of genetic mutations to help pinpoint the cause of diseases
https://deepmind.google/blog/alphamissense-catalogue-of-genetic-mutations-to-help-pinpoint-the-cause-of-diseases/

[2] EMBL-EBI: AlphaMissense data integrated into Ensembl, UniProt, DECIPHER and AlphaFold DB
https://www.ebi.ac.uk/about/news/technology-and-innovation/alphamissense-data-integration/

[3] ACGS: Best Practice Guidelines for Variant Classification in Rare Disease 2024
https://www.genomicseducation.hee.nhs.uk/wp-content/uploads/2024/08/ACGS-2024_UK-practice-guidelines-for-variant-classification.pdf

[4] Genetics in Medicine (2025): Calibration of additional computational tools expands ClinGen recommendation options for variant classification with PP3/BP4 criteria (author copy)
https://www.ccs.neu.edu/home/radivojac/papers/bergquist_genetmed_2025.pdf

[5] npj Genomic Medicine (2025): Discordance between a deep learning model and clinical-grade variant pathogenicity classification in a rare disease cohort
https://www.nature.com/articles/s41525-025-00480-w

Google DeepMind, with EMBL-EBI; calibrated for clinical use by the ClinGen Sequence Variant Interpretation Working Group website: https://deepmind.google

Version 1.0 · Published 20 November 2026 · Updated 20 November 2026 · Facts checked against sources 20 November 2026 · Downloaded 26 November 2026. This PDF may be out of date: the latest version, and a way to report an error or a better source, is at https://marluc.co.uk/case-studies/google-deepmind-alphamissense-variant-classification

VERSION HISTORY

v1.0 20 November 2026 First published