

time of the IMO 2024 competition, had not been defined in mathlib, AlphaProof struggled. To solve these problems, the authors relied on a specialized tool called Alpha-Geometry. A natural future avenue, then, is to target proofs that use custom definitions. This might require updates to the benchmarks used to evaluate and compare models (because current benchmarks typically assume fixed definitions that already exist), as well as further innovations in AI.

Despite its limitations, AlphaProof is the first AI tool I have used that I found concretely useful for writing proofs, a sentiment shared by some, but not all, of the mathematicians I have corresponded with. The lead author noted in his talk that there are plans to make the tool publicly available, although it is unclear to me how that would work given the financial and business constraints under which companies such as Google operate. In

any case, it is clear to me that something is changing in this field; perhaps AlphaProof is a harbinger of what is to come.

Talia Ringer is in the Siebel School of Computing and Data Science, University of Illinois at Urbana-Champaign, Urbana, Illinois, 61801, USA.
e-mail: tringer@illinois.edu

1. Hubert, T. *et al.* *Nature* **651**, 607–613 (2026).
2. *Nature* **644**, 20 (2025).
3. Wu, Y. *et al.* Preprint at arXiv <https://doi.org/10.48550/arXiv.2205.12615> (2022).
4. Ringer, T., Palmskog, K., Sergey, I., Gligoric, M. & Tatlock, Z. *Found. Trends Program. Lang.* **5**, 102–281 (2019).
5. Kontorovich, A. Preprint at arXiv <https://doi.org/10.48550/arXiv.2510.15924> (2025).

The author declares no competing interests.
This article was published online on 12 November 2025.

Medical research

AI succeeds in diagnosing rare diseases

Timo Lassmann

An artificial-intelligence system uses clinical data, genetic information and literature searches to suggest diagnoses and provides the underlying reasoning. **See p.775**

For the estimated 300 million people¹ living with a rare disease worldwide, obtaining an accurate diagnosis remains an arduous journey. The ‘diagnostic odyssey’ often takes five years or more², and is marked by repeated specialist consultations, misdiagnoses and unnecessary treatments. With more than 7,000 rare diseases recognized³, most of which affect fewer than one in 2,000 individuals, even experienced clinicians struggle to connect unfamiliar patterns of symptoms

with their underlying causes. On page 775, Zhao *et al.*⁴ present an artificial-intelligence system called DeepRare, which uses a range of specialized tools and knowledge sources to generate ranked diagnostic hypotheses for rare diseases, each accompanied by reasoning that links the conclusions to verifiable medical evidence.

The architecture of DeepRare reflects a broader shift in how AI systems tackle complex medical problems. Rather than relying

on a single model to produce answers, the system uses a coordinated approach. There is a central reasoning engine, powered by an AI system similar to those underlying chatbots such as ChatGPT, which coordinates multiple specialized steps, each responsible for distinct analytical tasks. One module extracts standardized symptom descriptions from clinical notes. Another searches the medical literature. A third retrieves similar cases from curated databases. Yet another analyses genetic variants when DNA-sequencing data are available. This division of labour mirrors the multidisciplinary team consultations that characterize assessments by specialists at rare-disease centres, but the AI tool operates at a scale and speed that is impossible for human teams.

The results are striking (Fig. 1). In a real-world clinical validation, DeepRare’s top-ranked prediction was correct 64% of the time, outperforming experienced rare-disease physicians who had access to search engines. When the correct diagnosis was among the system’s five ranked suggestions, accuracy reached 78.5%, compared with 65.6% for physicians. Perhaps most notably for clinical adoption, medical experts validated the system’s reasoning chains at a level of 95% agreement. This indicates that the explanations linking clinical data to diagnostic conclusions are clinically coherent and verifiable.

The system’s traceable reasoning might prove to be its most notable contribution to possible future use in the clinic. Rare-disease diagnosis is not merely about producing the correct answer; clinicians also need to understand why a diagnosis is plausible and what evidence supports it. DeepRare’s citation of specific sources, case reports, database entries and literature references provides an audit trail that supports, rather than supplants, clinical judgement. This transparency distinguishes the system from opaque AI approaches that might perform well but offer clinicians no basis for evaluating the recommendations made by AI. In high-stakes medical decisions, such interpretability is crucial.

What mechanisms underpin this high-performance AI? The authors systematically tested which components contribute most to accuracy, revealing two key factors. One is the retrieval of up-to-date medical knowledge from the web, and the other is a feature called self-reflection loops, which prompt the system to re-evaluate its preliminary conclusions. These capabilities prove particularly valuable given the rapidly evolving understanding of rare diseases, with hundreds of new genetic conditions described annually⁵. More generally, retrieval of current knowledge and iterative self-reflection are broadly applicable techniques that work across many different fields. Any field that requires reasoning using

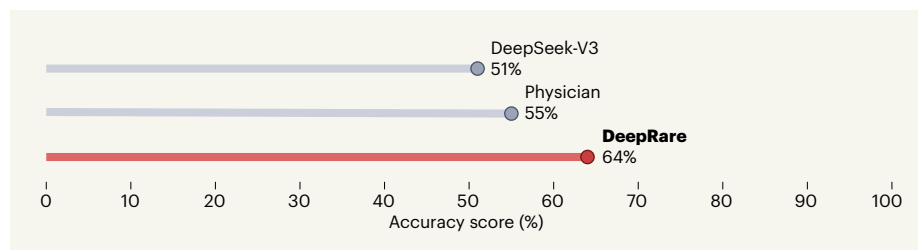


Figure 1 | Using artificial intelligence to diagnose rare diseases. Zhao *et al.*⁴ present a tool called DeepRare that can suggest five ranked diagnoses for rare diseases. The system uses a variety of information sources, including clinical data, DNA-sequencing results and searches of published papers. The results it provided are more accurate than those obtained by an early AI tool called DeepSeek-V3 or predictions made by physicians with access to a search engine. The score shown represents how often the system’s top prediction was accurate. (Adapted from Fig. 3c of ref. 4.)

up-to-date information could benefit from these AI techniques.

Several opportunities for refinement remain. A small, but non-negligible, fraction of generated reference citations were found to be ‘hallucinated’, meaning that the AI produced plausible-looking outputs that do not correspond to real sources. This is a known limitation of current AI systems that warrants attention before clinical deployment. This concern aligns with findings that large language models (LLMs) on their own recall only a fraction of the clinical features and gene associations held in specialist rare-disease databases, yet they can meaningfully complement existing database-driven diagnostic tools⁶. DeepRare tightly integrates LLMs with curated knowledge sources, rather than relying on LLMs alone.

DeepRare arrives amid the rapid expansion of AI capabilities across science, including advances such as autonomous chemical synthesis⁷. What distinguishes this AI system is its emphasis on collaboration rather than automation⁸. Showing the reasoning, not just the answer, is what makes AI useful when decisions carry real consequences. Whether in drug discovery, materials science or clinical diagnostics, the same principle holds: systems that show their workings enable scrutiny by human experts, earn trust and ultimately prove more useful than ‘black-box’ answers.

The diagnostic odyssey will not end with a single innovation. But AI that handles tedious legwork while remaining transparent about its sources and limitations is a meaningful step towards reducing the number of years that patients must wait for answers.

Timo Lassmann is at The Kids Research Institute Australia and the UWA Centre for Child Health Research, The University of Western Australia, Nedlands, Western Australia 6009, Australia.
e-mail: timo.lassmann@thekids.org.au

1. Nguengang Wakap, S. *et al.* *Eur. J. Hum. Genet.* **28**, 165–173 (2020).
2. Faye, F. *et al.* *Eur. J. Hum. Genet.* **32**, 1116–1126 (2024).
3. Haendel, M. *et al.* *Nature Rev. Drug Discovery* **19**, 77–78 (2019).
4. Zhao, W. *et al.* *Nature* **651**, 775–784 (2026).
5. Dawkins, H. J. S. *et al.* *Clin. Transl. Sci.* **11**, 11–20 (2018).
6. Groza, T. *et al.* *Hum. Genet. Genom. Adv.* **7**, 100558 (2026).
7. Boiko, D. A., MacKnight, R., Kline, B. & Gomes, G. *Nature* **624**, 570–578 (2023).
8. Zou, J. & Topol, E. J. *Lancet* **405**, 457 (2025).

The author declares no competing interests.
T.L. used Anthropic’s Claude Opus 4.5 to help prepare this News & Views article.
This article was published online on 18 February 2026.

Ecology

Rising temperatures pose a threat to tropical insects

Vojtech Novotny

An assessment of sensitivity to increasing temperature for thousands of insect species in mountainous terrain reveals a risk of insect biodiversity loss in tropical lowlands. **See p.672**

Recalling his sometimes fever-ridden travels in South America, the naturalist Alfred Russel Wallace wrote¹: “How many times, when almost overcome by the ague, had I crawled into the forest and been rewarded by some unknown and beautiful species!” He was not the only biologist fascinated by the diversity of insects in tropical forests. In a single lowland area of forest the number of insect species can exceed 25,000, according to current estimates². However, on page 672, Holzmänn *et al.*³ provide disconcerting evidence that this remarkable diversity could be threatened by climate change, particularly in the tropical lowlands.

The authors measured the maximum survival temperature, called the critical thermal maximum (CT_{max}), for 2,300 insect species along elevational gradients of mountainous terrain in Peru and Kenya, and found that in the lowlands, many species already live close to the maximum temperature they can tolerate. Global warming would render lowland forests uninhabitable for these species. This comprehensive evidence goes a long way towards overcoming scepticism about the prediction of widespread species loss in the lowlands, made in a pioneering biodiversity study from 2008 that examined an elevational gradient in Costa Rica⁴.

Global warming tends to accelerate with increasing elevation and latitude⁵, so both observed and predicted temperature increases remain relatively low for tropical lowlands. However, the stable temperature there means that even slow warming generates some of the world’s highest relative changes year to year (change-to-variability ratios), in which predicted annual temperature increase as a result of global warming exceeds its current year-to-year variation in average temperature⁵. It was suggested in the Costa Rica study⁴ that this could pose a danger to species adapted to a narrow range of stable temperatures, because these are about to change.

This proposition was not widely accepted at first because it was counter-intuitive, given that insect species in other ecosystems tolerate a wide range of temperatures.

However, there were limited data on physiological thermal limits for tropical species. Furthermore, forests in the tropics were already facing notable pressures from human activities such as logging, land-use conversion from forest to agriculture, and habitat fragmentation. Some biologists were therefore not ready to be alarmed by yet another, this time still theoretical, threat to biodiversity. However, CT_{max} measurements have gradually built a case for a threat to biodiversity in tropical lowlands⁶, paving the way for Holzmänn and colleagues’ study.

A close alignment of CT_{max} values with a stable ambient-temperature environment is expected, because maintaining a wide safety margin in terms of thermal tolerance could be physiologically costly. The flexibility of CT_{max}, whether through physiological responses in individual insects or rapid evolutionary changes, becomes crucial for the species’ ability to adapt to climate change. Holzmänn *et al.* tested lowland-insect species’ response to heat shocks and found limited flexibility to withstand such shocks.

Furthermore, using an innovative approach, the authors examined the thermal stability of 600 proteins from the studied species and found that the melting points at which the protein starts to unfold increased in tandem with CT_{max} values, further suggesting limited flexibility in thermal tolerance. The authors also found that CT_{max} does not change easily over the course of evolution, so closely related species usually have similar CT_{max} values. If the past tempo of CT_{max} evolution continues, species will not be able to evolve fast enough to cope with higher temperatures, given the contemporary rate of climate change⁷.

Insect diversity peaks either in the lowlands or at mid-elevations along elevational gradients, providing opportunities for species to relocate to cooler, higher terrain in response to warming, at least along long-distance elevational gradients⁴. However, the land area available at high elevations is limited in the tropics, especially outside South America⁸. This only increases the conservation value of