

AI-Driven Automation of Genetic Variant Interpretation: Enhancing Clinical Reporting with Large Language Models

Jean-Marie Ravel, Jiri Ruzicka, Nicolas Philippe, Dimitri Larue, Nicolas Duforet-Frebourg*, Jérôme Audoux*

* These authors contributed equally to this work

I. BACKGROUND | Why are Large Language Models (LLMs) perfectly suited for generating comments in genetic variant interpretation?



Time-consuming

Highly repetitive

Central part of the interpretation report

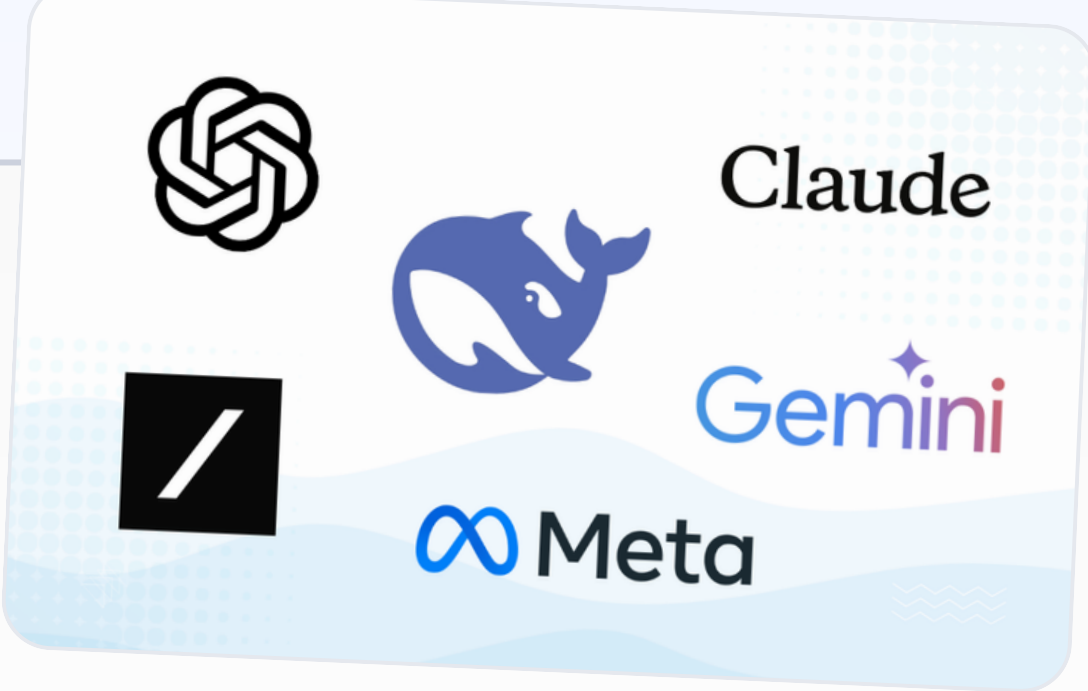
Bibliography: phenotype, therapeutic, etc.

Genetic variant interpretation is slow and complex
It requires integrating molecular data, clinical phenotypes, literature, and standardized rules such as ACMG guidelines.

Report requirements

Consistent structure

Brief ~~verbose~~



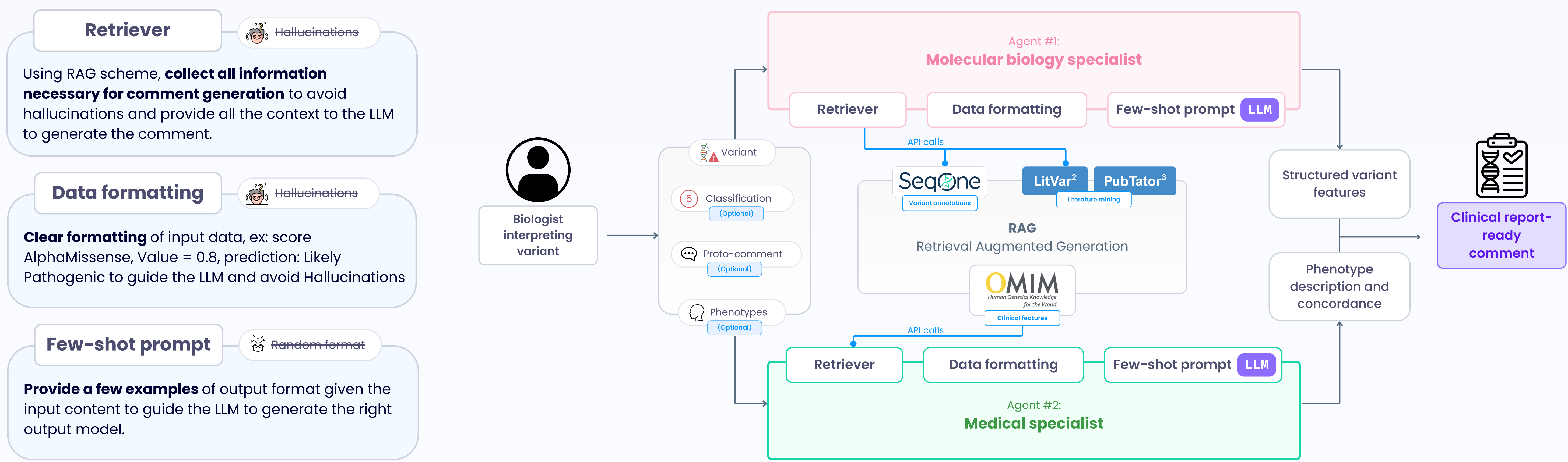
LLM limitations

Hallucinations

Random format

Large Language Models
can efficiently synthesize these diverse data sources to produce relevant, consistent, and reusable summaries. They are much faster than humans, but they need to be carefully integrated.

II. METHODS | AutoComment: Using LLM with multi-agents, RAG and few-shot prompting techniques.



III. RESULTS | Using LLM as a judge approach.

Classification

1 2 3 4 5

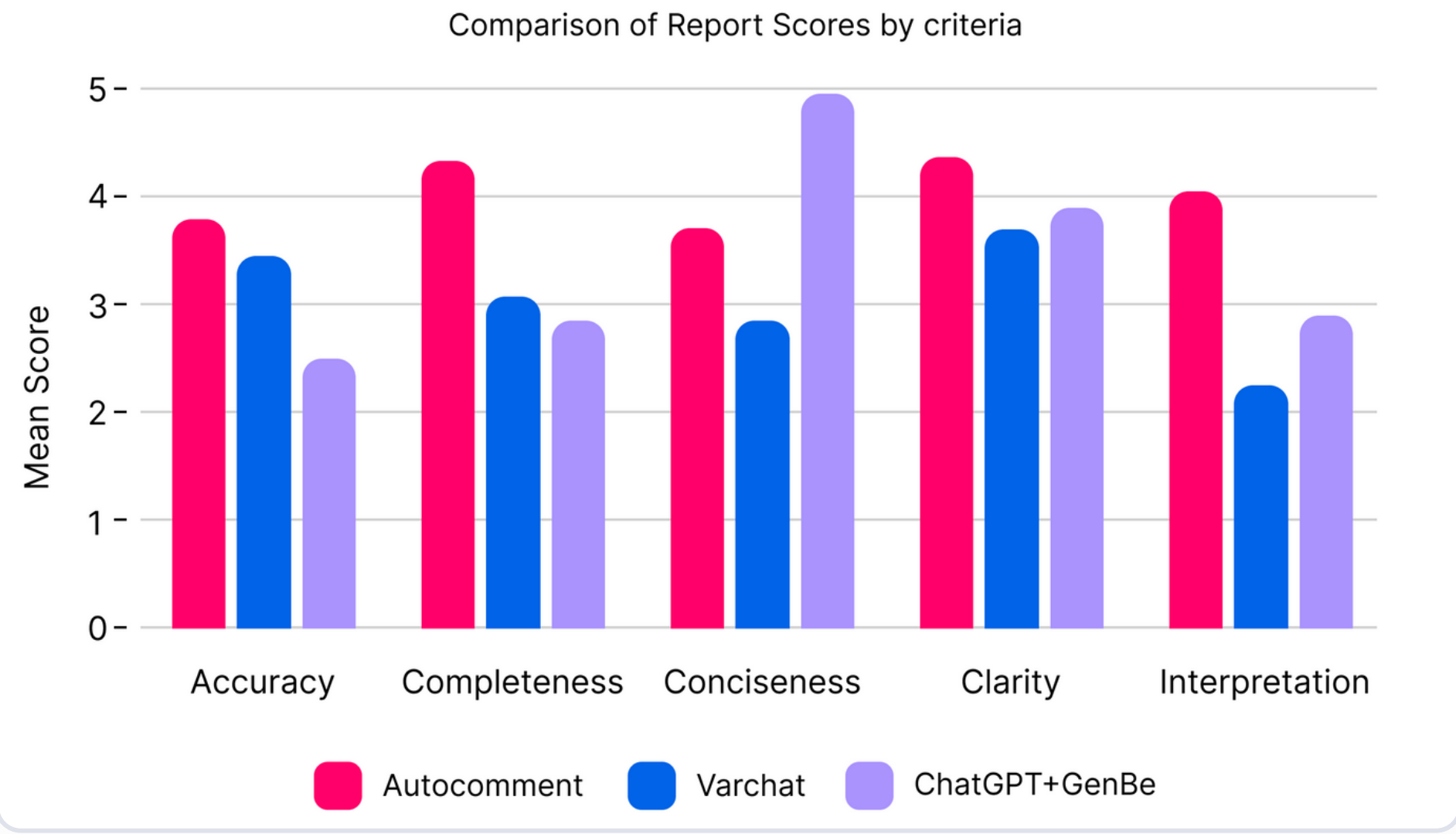
This is a missense variant located in exon 10 (10/11) of the ACADSB gene. The variant is rare, with a frequency of 0.0235% in gnomAD and 4 occurrences in the internal sample. In silico prediction tools (REVEL and CADD) support the pathogenic nature of the variation. The variant is classified as Pathogenic/Likely_pathogenic in ClinVar with multiple submissions and no conflicts. The missense variant is not predicted to cause loss of function and no impact on splicing is predicted by C-SpliceAI. The variant has been reported in literature (PMID: 36147814, 2022) in the context of Short/branched-chain acyl-CoA dehydrogenase deficiency (SBCADD), an inherited disorder of L-isoleucine metabolism.

Pathogenic variants in the ACADSB gene are associated with 2-methylbutyrylglycinuria, an autosomal recessive disorder. Clinical features vary widely, from asymptomatic individuals to severe cases with developmental delay, epilepsy, microcephaly, and eye abnormalities. Some reported symptoms like hypotonia and delayed motor development align with the provided HPO terms, but others such as ataxia, hydrocephalus, leukodystrophy, and optic neuropathy do not seem directly linked to this disease.

Reset comment Validate

Example of the interface
The interpretation was automatically generated by AutoComment, waiting for user modification / approval.

Comparison of Report Scores by criteria



Criteria	Autocomment	Varchat	ChatGPT+GenBe
Accuracy	3.8	3.4	2.5
Completeness	4.2	3.0	2.8
Conciseness	3.6	2.8	4.8
Clarity	4.2	3.6	3.8
Interpretation	4.0	2.2	2.8

LLM as a judge
We used Gemini 2.5 pro as an impartial evaluator to assess the quality of outputs, based on a given variant (n=21) generated by AutoComment (SeqOne), VarChat and ChatGPT with a direct input of GeneBe annotations.

IV. DISCUSSION

Consistent structure

Brief (verbose)

Hallucinations

Random-format

AutoComment is a trusted report generator
Using RAG and few-shot prompting, we present here a tool that improves biologists' efficiency. This work should be further validated and its implementation monitored.



DiagAI Copilot: Long term
This tool is one of many constituting the DiagAI tools suite that will be coordinate by a "controller agent", namely DiagAI Copilot.