A Bayesian model for quantifying genomic variant evidence sufficiency in Mendelian disease

Quant Group[1], Ali Saadat[2], and Dylan Lawless ** [3]

[1]The quantitative omics epidemiology group.
[2]Global Health Institute, School of Life Sciences, École Polytechnique Fédérale de Lausanne, Switzerland.
[3]Department of Intensive Care and Neonatology, University Children's Hospital Zurich, University of Zurich, Zurich, Switzerland.

December 13, 2025

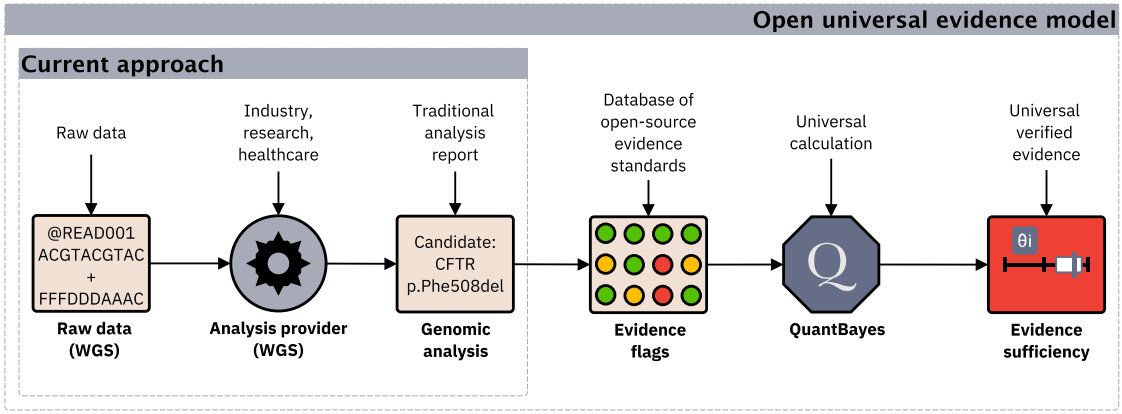
*Addresses for correspondence: Dylan.Lawless@kispi.uzh.ch

Abstract

Summary: Clinical genomic interpretation depends on heterogeneous evidence checks that vary across institutions and pipelines. This variation prevents evidence from being compared or verified in a reliable way. We address this logistical barrier by introducing a universal, verifiable layer that evaluates the completeness of available scientific evidence independently of upstream analytic systems. Quantitative Evidence Sufficiency (Quant ES) summarises, for each variant, how much verifiable evidence is available using a binary matrix derived from registered rule sets. A closed form Beta Binomial model produces a single evidence sufficiency estimate with a credible interval and genome wide percentile. This separates evidence availability from provider-specific logic and enables consistent interpretation and reuse of results across institutions.

Implementation: QuantBayes implements the method as compiled C binaries and an R package. It requires no access to proprietary algorithms and scales to millions of variants, allowing integration into existing diagnostic and research workflows without modification to upstream systems.

Availability: Source code and data are available at the Zenodo repository (pending) and software releases at https://doi.org/10.5281/zenodo.17919369 (1). Released under the MIT licence.



Graphical abstract. Genomic analyses produce candidate variants, but the completeness of their supporting evidence is not directly comparable across institutions. Quant Evidence Sufficiency operates on a binary evidence matrix to summarise how much verifiable scientific evidence is available for each variant. This separates analysis outputs from interpretation and delivers an interpretable, uncertainty-calibrated measure of credibility that can be shared and reused without exposing provider-specific methods or intellectual property.

1 Introduction

Clinical variant interpretation relies on analysis that include biological plausibility, inheritance consistency, population frequency, and phenotype fit (2–7). These assessments are implemented differently across pipelines, so the evidential basis for a candidate cannot be reliably compared or verified across institutions (8). Structured specifications such as the qualifying variant framework (9) and the American College of Medical Genetics and Genomics (ACMG) criteria (2; 3) provide practical ways to generate verifiable evidence flags. These include checks such as confirming inheritance consistency or excluding variants that are too common to explain a rare disorder. These rules can be evaluated for any pipeline result without disclosure of proprietary logic.

Nevertheless, the field lacks a statistical measure of how complete and trustworthy the resulting verifiable evidence is for a specific candidate. Manual inspection and heuristic scoring do not quantify uncertainty and do not place candidates within the wider evidential landscape.

This gap reflects a broader logistical challenge. Institutions encode and interpret evidence differently, yet clinical decision making requires a shared, reproducible basis for assessing how well supported each candidate is. A framework is needed that operates solely on verifiable evidence, remains independent of pipeline internals, and quantifies evidence completeness in a transparent and comparable manner.

Here we address this challenge by introducing Quantitative Evidence Sufficiency model (Quant ES), a Bayesian framework (10; 11) that models the sufficiency of verifiable evidence for each variant from its binary evidence profile. The approach uses a conjugate Beta Binomial formulation to express uncertainty and to situate each variant within the wider evidential landscape without reliance on provider specific algorithms. This framework introduces a distinct evidential layer that complements, rather than replaces, pathogenicity prediction. Existing tools typically estimate the likelihood that a variant is causal, whereas the completeness and reliability of the underlying scientific evidence are rarely measured directly (12). Quant ES fills this gap by quantifying how much verifiable information is available for each candidate. By providing a universal evidential measure applicable across clinical, research, and commercial settings, Quant ES supports transparent interpretation and strengthens interoperability across institutions.

2 Methods

2.1 Evidence matrix

Genomic variant calling, annotation, and interpretation are implemented heterogeneously across institutions, limiting direct comparability of variant reports (2–7; 12–15). To demonstrate a verifiable and comparable input, we briefly summarise an ex-

ternally defined and registered qualifying variant evidence standard (8; 9). This specification is independent of Quant ES and encodes rule-based evidence derived from subject-level data and reference databases into standardised evidence flags, which are subsequently interpreted as evidence present or missing (examples in **Box S1–S2** and **Table S1**). QuantES operates solely on the resulting binary evidence matrix, irrespective of how the flags are generated. For each variant i, the interpretation rule set (8; 9) produces a binary vector of length m, where

$$X_{ij} = \begin{cases} 1 & \text{if evidence rule } j \text{ is satisfied,} \\ 0 & \text{otherwise.} \end{cases}$$

The total number of satisfied rules is $k_i = \sum_{j=1}^{m} X_{ij}$, and the empirical proportion of satisfied evidence is $\hat{p}_i = k_i/m$. This quantity summarises the amount of verifiable evidence supporting variant i.

Rules are defined so that a TRUE flag always indicates a contradiction or weakening signal and is encoded as $X_{ij} = 0$. A FALSE flag indicates that no contradiction is observed and is encoded as $X_{ij} = 1$. Missing or unavailable information is also encoded as $X_{ij} = 0$ and never contributes evidence. For example, in **Table S1**, a TRUE flag for parental genotype availability indicates that the genotype is unavailable and contributes no evidence ($X_{ij} = 0$), whereas a FALSE flag indicates that the genotype is available and contributes evidence ($X_{ij} = 1$).

Bayesian evidence sufficiency model

We define the evidence sufficiency parameter θ_i as the probability that a randomly selected evidence rule is satisfied for variant i. The model assumes $k_i \mid \theta_i \sim \text{Binomial}(m, \theta_i), \quad \theta_i \sim \text{Beta}(1, 1)$. By conjugacy, the posterior distribution is $\theta_i \mid k_i \sim \text{Beta}(1+k_i,\ 1+m-k_i)$. The posterior mean and Credible Interval (CrI) (11; 16) provide an uncertainty-calibrated measure of evidence completeness for each variant (10; 17).

Global posterior distribution

To contextualise variant-level evidence, we analyse each θ_i relative to the empirical mixture of all posterior distributions across the genome (18). Monte Carlo draws from each θ_i yield a global distribution from which we estimate genome-wide summary statistics and CrI. For a set of candidate variants, we compute the percentile of each θ_i within this distribution as $\text{percentile}(i) = \Pr\left(\theta_j \leq \theta_i\right)$, with j indexing all evaluated variants. This provides a genome-wide reference for interpreting evidence sufficiency.

Validation study on a WGS trio

We evaluated the framework using Whole Genome Sequencing (WGS) from the public Genome in a Bottle (GIAB) family trio (HG005–HG007, PRJNA200694, GRCh38 v4.2.1) (19). All processing parameters were defined in an external qualifying variant specification (9). Quality control and study filters were applied to illustrate a typical clinical WGS pipeline. Site-level thresholds on QUAL and INFO/DP were applied using BCFtools, complemented by per-sample thresholds on FORMAT/DP and FORMAT/GQ with exclusion of missing genotypes. Composite filters required either all trio samples or at least one sample to pass the criteria. The filtered trio Variant Call Format (VCF) was analysed with Exomiser (default configuration) using the family `.ped` file and without Human Phenotype Ontology (HPO) terms (20). Exomiser outputs were then processed with the registered qualifying variant evidence standard (8; 9) flag rule set to generate raw per-variant evidence flags (TRUE, FALSE, or NA). A second registered interpretation rule set converted these raw flags into the binary evidence matrix required by Quant ES, with values indicating evidence supporting (1) or missing (0). All rule sets used in this workflow are provided in YAML Ain't Markup Language (YAML) or JavaScript Object Notation (JSON) format (9). Quant ES was then run on the binary evidence matrix using default settings to generate per-variant verifiable evidence θ, 95% CrIs, and genome-wide evidence summaries, implemented with `QuantBayes` software.

3 Implementation

QuantBayes software

`QuantBayes` implements the Quant ES model as a lightweight, pipeline-independent tool. It is distributed as precompiled C binaries for macOS and Linux and as a cross platform R package with identical statistical behaviour (1). All implementations operate directly on a binary evidence matrix containing values in {0, 1, NA} and require no access to genotype data, pipeline internals, or proprietary algorithms. The expected input format is illustrated in **Figure 1 (A)**. The standalone C binary runs without installation and is suitable for automated High Performance Computer (HPC) workflows. It produces a per-variant summary including the number of satisfied evidence rules k_i, the number of evaluated rules m, the posterior mean of θ_i, its 95% CrI, and the percentile within the genome wide distribution, together with a global summary of the posterior mixture. Optional outputs include a concise human-readable report and structured JSON representations for integration into relational or graph based systems such as Electronic Health Record (EHR) platforms or Resource Description Framework (RDF) stores (27). Per-variant outputs are shown in **Figure 1 (B)**, with an overview of computed quantities in **Figure S2**. All implementations use the same closed form Beta updates and yield identical results up to floating point precision.

Reproducibility and compatibility

Input and output formats are compatible with standard variant tables, including decomposed VCF records and PLINK text filesets (21; 22). Variant identifiers may follow the VCF `CHROM-POS-REF-ALT` convention when the `ID` field is absent, and outputs preserve input row-order. All computations in Quant ES use deterministic evaluation of the Beta posterior. Where Monte Carlo sampling is used for genome-wide summaries, fixed seeds may be supplied. Evidence standards and interpretation rule sets are provided as registered YAML or JSON specifications (9), ensuring consistent reconstruction of evidence matrices and posterior outputs across sites. Integration

with workflow managers such as Snakemake (23) and Nextflow (24) supports reproducible deployment in existing analysis environments.

4 Results

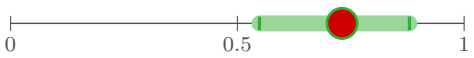
Per-variant summaries

Figure 1 (B) shows a basic `QuantBayes` output. A structured report generated from the same values is provided in **Box 1**. Traditionally, variant interpretation pipelines identify one or more candidate causal variants based on their internal analyses. `QuantBayes` answers a different question: it quantifies, for every variant, how much verifiable internal and external scientific evidence is available, so the strength of support for any nominated candidate is reported. Using the rule set that defines the evidence flags, `QuantBayes` provides this transparent and reproducible measure that can be understood and compared across institutions.

Box 1: Example of verifiable evidence sufficiency output

QV-ES: `qv_variant_interpretation.yaml`
QV-ES registered MD5: 6343b82a254edc709f4dc766fe056253

Case: autosomal recessive cystic fibrosis with a primary candidate variant in *CFTR*.
Candidate variant: `NM_000492.4:c.1521_1523delCTT (p.Phe508del)`.
Percentile: 99.875
Evidence sufficiency θ (95% CrI): 0.7308 (0.5487 – 0.8793)
Evidence scale:



Total evaluated variants: 400
Total evaluated evidence features (m): 24
Evidence rules satisfied for this variant (k): 18
Global evidence sufficiency θ 95% CrI: 0.475 (0.082 – 0.951).
Higher values indicate that most required evidence elements are available; lower values reflect sparse or incomplete support. For this candidate, many required domains are satisfied. Across all variants, the mean θ is 0.475 and the median is 0.459, reflecting the typical level of interpretable evidence in this case.

Validation study evaluation on a WGS trio

The Quant ES framework was applied to the GIAB Trio (HG005–HG007, GRCh38 v4.2.1) and analysed with Exomiser (20), producing a ranked annotated variant list typical of clinical WGS analysis for Mendelian disease. **Figure S7 - S8** show the evidence standard flag rules before and after interpretation, reflecting the confirmation of evidence features such as population frequency and inheritance consistency.

Figure 1 (C) displays the genome-wide mixture distribution of θ with overlaid density curves for the top Exomiser candidates. The variant with the strongest verifiable evidence sufficiency was *RYR2*:ENST00000366574.7:c.6493dup:p.(Leu2165Profs*44) with Autosomal Dominant (AD) Mode of Inheritance (MOI), with θ (95% CrI) =

A

Input

Variant ID	rule1	rule2	rule3	...	rule *m*
7-117559592-CTT-N	1	0	1	...	1
6-72183475-CG-N	0	1	1	...	0
1-14682421-G-A	1	1	0	...	1

B

Output

Variant ID	*k*	*m*	*θ* mean	*θ* lower	*θ* upper	percentile
7-117559592-CTT-N	18.00	24.00	0.73	0.55	0.88	99.88
6-72183475-CG-N	18.00	24.00	0.73	0.55	0.88	99.88
1-14682421-G-A	17.00	24.00	0.69	0.51	0.85	99.38
...						
Global	24.00	24.00	0.52	0.54	0.38	0.65

C

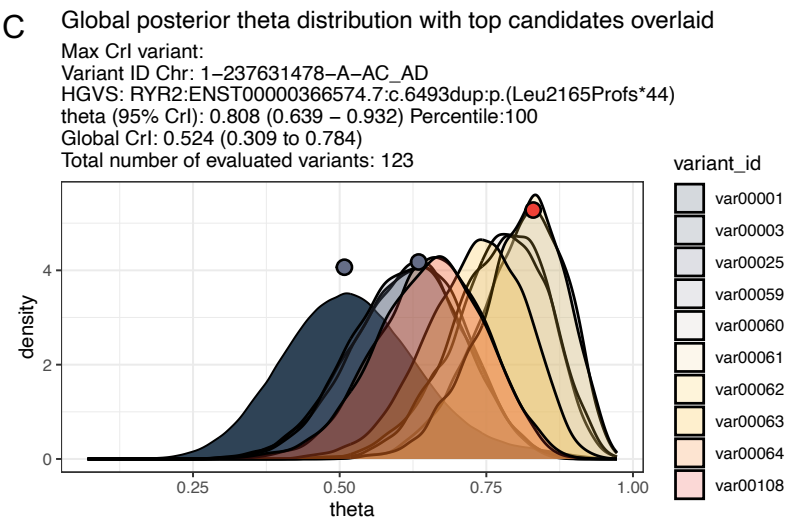


Figure 1: **Input, and output, and GIAB WGS trio study.** **(A)** Binary evidence matrix with variants as rows and evidence rules as columns, indicating whether each rule is satisfied (1) or not (0). **(B)** QuantBayes output summarising, per variant, the number of satisfied rules k, total evaluated rules m, posterior mean θ, its 95% CrI, and the genome-wide percentile of θ. **(C)** Genome-wide mixture distribution of θ across all variants (dark), with transparent overlays for top Exomiser-prioritised candidates. The red dot marks the variant with the highest evidence sufficiency CrI, and grey dots indicate the two phenotype-based Exomiser rank 1 predictions. Elevated values reflect stronger interpretative evidence, not causality.

0.808 (0.639 – 0.932) and a percentile of 100. This candidate emerged from 123 evaluated variants derived from an initial WGS call set of 5,188,814 variants. Variants rising above the genome-wide mixture are supported by substantially more interpretable evidence. Quant ES therefore quantifies the strength of support for hypotheses proposed by tools such as Exomiser, in a manner directly comparable across approximately one hundred variant effect and prioritisation predictors (12; 20; 25).

Taken together, these results show that Quant ES provides a reproducible, transparent, pipeline independent measure of evidence availability. The framework separates the evidence layer from provider-specific logic and supports consistent interpretation, verification, and shared use across clinical, research, and commercial settings.

Performance benchmarks

`QuantBayes` runs efficiently across both clinical and cohort-scale analyses (Figure S1). Runtime scales with the number of variants and evaluated evidence rules. Clinical diagnostic-sized inputs complete in negligible time, while matrices containing up to ten million variants and five hundred evidence sources complete in under ten minutes on a standard laptop. Precompiled C binaries run directly on HPC systems without requiring administrator privileges. The method is therefore practical for routine clinical use and large cohort studies without specialised hardware.

Public and patient involvement

Public and Patient Involvement and Engagement (PPIE) in genomics is constrained by a persistent practical problem where patients and secondary institutions cannot guarantee which analyses were performed, and results therefore cannot be reused without repeating analysis (26). This arises because evidence checks are embedded within local pipelines rather than recorded in a standard portable, verifiable form. QuantBayes resolves this by running on a pipeline-independent checklist of verifiable evidence that accompanies genomic results. Evidence availability is encoded as a registered, versioned binary matrix derived from explicit rule sets, making clear which

evidence domains were evaluated and which were satisfied for each variant. Building on the qualifying variant framework (9), which captured analytic intent and PPIE input as structured metadata, QuantBayes extends this principle by enabling PPIE groups to contribute to the construction of standardised evidence rule sets, supporting transparent verification and reuse of genomic results across institutions (27; 28).

5 Conclusion

We introduced a Bayesian framework for quantifying how complete the verifiable scientific evidence is for each genomic variant in Mendelian disease analysis. By operating directly on an evidence matrix and remaining independent of provider specific pipelines, the method delivers a transparent and reproducible evidential layer. A single posterior sufficiency parameter, with uncertainty and genome wide context, summarises both variant level evidence completeness and relative standing among alternatives. This enables consistent interpretation, direct comparison of results across institutions, and reliable reuse of genomic evidence without reanalysis, supporting interoperability in genomic medicine.

Funding

This project was supported through the grant Swiss National Science Foundation 320030_201060, and NDS-2021-911 (SwissPedHealth) from the Swiss Personalized Health Network and the Strategic Focal Area 'Personalized Health and Related Technologies' of the ETH Domain (Swiss Federal Institutes of Technology).

Acknowledgements

Acknowledgements We would like to thank all the patients and families who have been providing advice on SwissPedHealth and its projects, as well as the clinical and research teams at the participating institutions.

Data and software availability

All data used in this study are derived from open sources referenced in the Methods section. WGS GIAB trio data (HG005–HG007, PRJNA200694, GRCh38 v4.2.1) described by Wagner et al. (19), together with the benchmark small variant VCFs on GRCh38 available at https://ftp-trace.ncbi.nlm.nih.gov/ReferenceSamples/giab/release/. All results are reproducible using the accompanying code repository. Generated data are deposited in the Zenodo archive. The Zenodo archive includes precompiled `QuantBayes` binaries for macOS (Intel), Linux (x86_64), and R package, https://doi.org/10.5281/zenodo.17919369 (1). The `QuantBayes` R package source code is also available at https://github.com/DylanLawless/src-quantbayes_package. A CRAN release is planned. All software is provided under MIT license.

Contributions

DL produced the work and contributed to manuscript writing. AS contributed to the manuscript writing and software testing. LJS contributed to the manuscript writing and funding.

Competing interests

The authors declare no competing interests.

Ethics statement

The projects were approved by the respective ethics committees of all participating centers (Cantonal Ethics Committee Bern, approval number KEK-029/11) and the study was conducted in accordance with the Declaration of Helsinki.

Acronyms

ACMG American College of Medical Genetics and Genomics 3

AD Autosomal Dominant 9

CrI Credible Interval 5

EHR Electronic Health Record 7

GIAB Genome in a Bottle 6

HPC High Performance Computer 7

HPO Human Phenotype Ontology 6

JSON JavaScript Object Notation 6

MOI Mode of Inheritance 9

PPIE Public and Patient Involvement and Engagement 11

Quant ES Quantitative Evidence Sufficiency model 4

RDF Resource Description Framework 7

VCF Variant Call Format 6

WGS Whole Genome Sequencing 6

YAML YAML Ain't Markup Language 6

References

[1] Dylan Lawless. Bayesian quantification of evidence sufficiency quantbayes implementation. December 2025. doi: 10.5281/zenodo.17919369. URL https://doi.org/10.5281/zenodo.17919369.

[2] Sue Richards, Nazneen Aziz, Sherri Bale, David Bick, Soma Das, Julie Gastier-Foster, Wayne W Grody, Madhuri Hegde, Elaine Lyon, Elaine Spector, et al.

Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the american college of medical genetics and genomics and the association for molecular pathology. *Genetics in medicine*, 17 (5):405–423, 2015.

[3] Marilyn M Li, Michael Datto, Eric J Duncavage, Shashikant Kulkarni, Neal I Lindeman, Somak Roy, Apostolia M Tsimberidou, Cindy L Vnencak-Jones, Daynna J Wolff, Anas Younes, et al. Standards and guidelines for the interpretation and reporting of sequence variants in cancer: a joint consensus recommendation of the association for molecular pathology, american society of clinical oncology, and college of american pathologists. *The Journal of molecular diagnostics*, 19(1):4–23, 2017.

[4] Quan Li and Kai Wang. Intervar: clinical interpretation of genetic variants by the 2015 acmg-amp guidelines. *The American Journal of Human Genetics*, 100 (2):267–280, 2017.

[5] Erin Rooney Riggs, Erica F Andersen, Athena M Cherry, Sibel Kantarci, Hutton Kearney, Ankita Patel, Gordana Raca, Deborah I Ritter, Sarah T South, Erik C Thorland, et al. Technical standards for the interpretation and reporting of constitutional copy-number variants: a joint consensus recommendation of the american college of medical genetics and genomics (acmge and the clinical genome resource (clingen). *Genetics in Medicine*, 22(2):245–257, 2020.

[6] Sean V Tavtigian, Steven M Harrison, Kenneth M Boucher, and Leslie G Biesecker. Fitting a naturally scaled point system to the acmg/amp variant classification guidelines. *Human mutation*, 41(10):1734–1737, 2020.

[7] Brent S Pedersen, Joe M Brown, Harriet Dashnow, Amelia D Wallace, Matt Velinder, Martin Tristani-Firouzi, Joshua D Schiffman, Tatiana Tvrdik, Rong Mao, D Hunter Best, et al. Effective variant filtering and expected candidate variant yield in studies of rare human disease. *NPJ Genomic Medicine*, 6(1): 1–8, 2021.

[8] Swiss Genomics Association. Swiss genomics association consensus guideline for evidence based genomic variant interpretation in mendelian disease. 2025. URL https://www.swissgenomicsassociation.ch/releases/. Draft in progress. Final publication anticipated.

[9] Dylan Lawless, Ali Saadat, Mariam Ait Oumelloul, Simon Boutry, Veronika Stadler, Sabine Österle, Jan Armida, David Haerry, D. Sean Froese, Luregn J. Schlapbach, and Jacques Fellay. Application of qualifying variants for genomic analysis. *medrxiv*, May 2025. doi: 10.1101/2025.05.09.25324975.

[10] José M. Bernardo and Adrian F. M. Smith. *Bayesian Theory*. Wiley Series in Probability and Statistics. Wiley, 1 edition, May 1994. ISBN 978-0-471-49464-5 978-0-470-31687-0. doi: 10.1002/9780470316870.

[11] Andrew Gelman, John B. Carlin, Hal S. Stern, and Donald B. Rubin. *Bayesian Data Analysis*. Chapman and Hall/CRC, 0 edition, June 1995. ISBN 978-1-135-43941-5. doi: 10.1201/9780429258411.

[12] Cristian Riccio, Max L. Jansen, Linlin Guo, and Andreas Ziegler. Variant effect predictors: A systematic review and practical guide. *Human Genetics*, 143(5): 625–634, May 2024. ISSN 1432-1203. doi: 10.1007/s00439-024-02670-5.

[13] Carl A Anderson, Fredrik H Pettersson, Geraldine M Clarke, Lon R Cardon, Andrew P Morris, and Krina T Zondervan. Data quality control in genetic case-control association studies. *Nature protocols*, 5(9):1564–1573, 2010. URL https://doi.org/10.1038/nprot.2010.116.

[14] Emil Uffelmann, Qin Qin Huang, Nchangwi Syntia Munung, Jantina De Vries, Yukinori Okada, Alicia R Martin, Hilary C Martin, Tuuli Lappalainen, and Danielle Posthuma. Genome-wide association studies. *Nature Reviews Methods Primers*, 1(1):59, 2021. URL https://doi.org/10.1038/s43586-021-00056-9.

[15] Brent S Pedersen and Aaron R Quinlan. Vcfexpress: flexible, rapid user-expressions to filter and format VCFs. *Bioinformatics*, 41(3):

btaf097, March 2025. ISSN 1367-4811. doi: 10.1093/bioinformatics/btaf097. URL https://academic.oup.com/bioinformatics/article/doi/10.1093/bioinformatics/btaf097/8051444.

[16] E. T. Jaynes. *Probability Theory: The Logic of Science*. Cambridge University Press, 1 edition, April 2003. ISBN 978-0-521-59271-0 978-0-511-79042-3. doi: 10.1017/CBO9780511790423.

[17] Norman L. Johnson, Adrienne W. Kemp, and Samuel Kotz. *Univariate Discrete Distributions*. Wiley Series in Probability and Statistics. Wiley, 1 edition, August 2005. ISBN 978-0-471-27246-5 978-0-471-71581-8. doi: 10.1002/0471715816.

[18] Bradley Efron. *Large-Scale Inference: Empirical Bayes Methods for Estimation, Testing, and Prediction*. Cambridge University Press, 1 edition, August 2010. ISBN 978-0-521-19249-1 978-0-511-76136-2 978-1-107-61967-8. doi: 10.1017/CBO9780511761362.

[19] Justin Wagner, Nathan D. Olson, Lindsay Harris, Ziad Khan, Jesse Farek, Medhat Mahmoud, Ana Stankovic, Vladimir Kovacevic, Byunggil Yoo, Neil Miller, Jeffrey A. Rosenfeld, Bohan Ni, Samantha Zarate, Melanie Kirsche, Sergey Aganezov, Michael C. Schatz, Giuseppe Narzisi, Marta Byrska-Bishop, Wayne Clarke, Uday S. Evani, Charles Markello, Kishwar Shafin, Xin Zhou, Arend Sidow, Vikas Bansal, Peter Ebert, Tobias Marschall, Peter Lansdorp, Vincent Hanlon, Carl-Adam Mattsson, Alvaro Martinez Barrio, Ian T. Fiddes, Chunlin Xiao, Arkarachai Fungtammasan, Chen-Shan Chin, Aaron M. Wenger, William J. Rowell, Fritz J. Sedlazeck, Andrew Carroll, Marc Salit, and Justin M. Zook. Benchmarking challenging small variants with linked and long reads. *Cell Genomics*, 2(5):100128, May 2022. ISSN 2666979X. doi: 10.1016/j.xgen.2022.100128.

[20] Valentina Cipriani, Nikolas Pontikos, Gavin Arno, Panagiotis I. Sergouniotis, Eva Lenassi, Penpitcha Thawong, Daniel Danis, Michel Michaelides, Andrew R. Webster, Anthony T. Moore, Peter N. Robinson, Julius O.B. Jacobsen, and

Damian Smedley. An Improved Phenotype-Driven Tool for Rare Mendelian Variant Prioritization: Benchmarking Exomiser on Real Patient Whole-Exome Data. *Genes*, 11(4):460, April 2020. ISSN 2073-4425. doi: 10.3390/genes11040460.

[21] Shaun Purcell, Benjamin Neale, Kathe Todd-Brown, Lori Thomas, Manuel A.R. Ferreira, David Bender, Julian Maller, Pamela Sklar, Paul I.W. De Bakker, Mark J. Daly, and Pak C. Sham. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics*, 81(3):559–575, September 2007. ISSN 00029297. doi: 10.1086/519795.

[22] Petr Danecek, Adam Auton, Goncalo Abecasis, Cornelis A. Albers, Eric Banks, Mark A. DePristo, Robert E. Handsaker, Gerton Lunter, Gabor T. Marth, Stephen T. Sherry, Gilean McVean, Richard Durbin, and 1000 Genomes Project Analysis Group. The variant call format and VCFtools. *Bioinformatics*, 27(15):2156–2158, August 2011. ISSN 1367-4811, 1367-4803. doi: 10.1093/bioinformatics/btr330.

[23] Felix Mölder, Kim Philipp Jablonski, Brice Letcher, Michael B. Hall, Christopher H. Tomkins-Tinch, Vanessa Sochat, Jan Forster, Soohyun Lee, Sven O. Twardziok, Alexander Kanitz, Andreas Wilm, Manuel Holtgrewe, Sven Rahmann, Sven Nahnsen, and Johannes Köster. Sustainable data analysis with Snakemake. *F1000Research*, 10:33, January 2021. ISSN 2046-1402. doi: 10.12688/f1000research.29032.1. URL https://f1000research.com/articles/10-33/v1.

[24] Paolo Di Tommaso, Maria Chatzou, Evan W Floden, Pablo Prieto Barja, Emilio Palumbo, and Cedric Notredame. Nextflow enables reproducible computational workflows. *Nature Biotechnology*, 35(4):316–319, April 2017. ISSN 1087-0156, 1546-1696. doi: 10.1038/nbt.3820. URL https://www.nature.com/articles/nbt.3820.

[25] Manuel Holtgrewe, Oliver Stolpe, Mikko Nieminen, Stefan Mundlos, Alexej Knaus, Uwe Kornak, Dominik Seelow, Lara Segebrecht, Malte Spielmann, Björn Fischer-Zirnsak, Felix Boschann, Ute Scholl, Nadja Ehmke, and Dieter Beule.

VarFish: Comprehensive DNA variant analysis for diagnostics and research. *Nucleic Acids Research*, 48(W1):W162–W169, July 2020. ISSN 0305-1048, 1362-4962. doi: 10.1093/nar/gkaa241.

[26] Mark D Wilkinson, Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino da Silva Santos, Philip E Bourne, et al. The fair guiding principles for scientific data management and stewardship. *Scientific data*, 3(1):1–9, 2016.

[27] Vasundra Touré, Philip Krauss, Kristin Gnodtke, Jascha Buchhorn, Deepak Unni, Petar Horki, Jean Louis Raisaro, Katie Kalt, Daniel Teixeira, Katrin Crameri, et al. Fairification of health-related data using semantic web technologies in the swiss personalized health network. *Scientific Data*, 10(1):127, 2023.

[28] Eelke van der Horst, Deepak Unni, Femke Kopmels, Jan Armida, Vasundra Touré, Wouter Franke, Katrin Crameri, Elisa Cirillo, and Sabine Österle. Bridging clinical and genomic knowledge: An extension of the sphn rdf schema for seamless integration and fairification of omics data. 2023.

[29] Robert E. Kass and Larry Wasserman. The selection of prior distributions by formal rules. *Journal of the American statistical Association*, 91(435):1343–1370, 1996.

[30] A. P. Dawid. Causal Inference without Counterfactuals. *Journal of the American Statistical Association*, 95(450):407–424, June 2000. ISSN 0162-1459, 1537-274X. doi: 10.1080/01621459.2000.10474210.

[31] Judea Pearl. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, 2 edition, September 2009. ISBN 978-0-511-80316-1 978-0-521-89560-6 978-0-521-74919-0. doi: 10.1017/CBO9780511803161.

[32] Michael S. Hamada, Alyson G. Wilson, C. Shane Reese, and Harry F. Martz. *Bayesian Reliability*. Springer Series in Statistics. Springer New York, New York, NY, 2008. ISBN 978-0-387-77948-5 978-0-387-77950-8. doi: 10.1007/978-0-387-77950-8.

Supplemental

A Bayesian model for quantifying genomic variant evidence sufficiency in Mendelian disease

Example QV ES flags and interpretation rules

Box S1: Example QV ES YAML (parental genotype completeness)

```yaml
criteria:
  parent_gt_unavailable_mother:
    conditions:
      - field: mother_geno
        operator: '=='
        value: NA
    description: Maternal genotype is unavailable
    logic: and
  parent_gt_unavailable_any:
    conditions:
      - field: mother_geno
        operator: '=='
        value: NA
      - field: father_geno
        operator: '=='
        value: NA
    description: At least one parental genotype is unavailable
    logic: or
notes:
  - parent_gt_unavailable_mother="The maternal genotype is unavailable.
  Any inheritance or segregation checks requiring maternal information
  cannot be evaluated, reducing the completeness of verifiable evidence
  for this variant."
  - parent_gt_unavailable_any="At least one parental genotype is
  unavailable. Inheritance and family-based consistency checks relying
  on parental information cannot be fully evaluated, reducing the
  completeness of verifiable evidence for this variant."
```

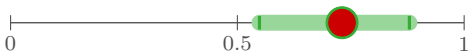
Box S2: Example QV ES YAML (gnomAD quality flags)

```
criteria:
  gnomad_flag_present:
    conditions:
      - field: gnomad_flag
        operator: in
        value: AC0,AS-VQSR,InbreedingCoeff,RF,Discrepant
        Frequencies,CHIP,Monoallelic,Only heterozygous,MNV,LCR,
        LC pLoF,pLoF Flag,NC Transcript,SEGDUP,Common low
        heteroplasmy
    description: At least one gnomAD quality or annotation flag is
    present
    logic: or
notes:
  - 'gnomad_flag_present= gnomAD flags are retained for review;
  presence of any flag requires explicit mention or justification
  in final causal report. For information see:
  https://gnomad.broadinstitute.org/help'
```

Box S3: Example of verifiable evidence sufficiency output

QV-ES: `qv_variant_interpretation.yaml`
QV-ES registered MD5: 6343b82a254edc709f4dc766fe056253

Clinical scenario: autosomal recessive cystic fibrosis with a primary candidate variant in *CFTR*.
Candidate variant: `NM_000492.4:c.1521_1523delCTT (p.Phe508del)`.
Percentile: 99.875
Evidence sufficiency θ (95% CrI): 0.7308 (0.5487 – 0.8793)
Evidence scale:



Further information
Total evaluated variants: 400
Total evaluated evidence features (m): 24
Evidence rules satisfied for this variant (k): 18
Global evidence sufficiency θ 95% CrI: 0.475 (0.082 – 0.951).
Higher values indicate that most required evidence elements are available; lower values reflect sparse or incomplete support. For this candidate, many required domains are satisfied. Across all variants, the mean θ is 0.475 and the median is 0.459, reflecting the typical level of interpretable evidence in this case.

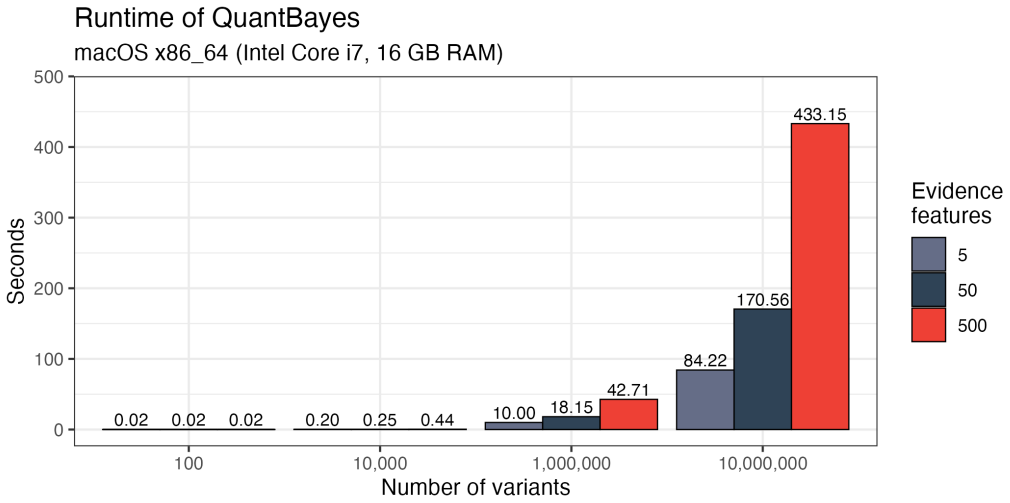


Figure S1: Runtime of QuantBayes for varying numbers of variants (rows) and evidence features (columns). Clinical sized inputs complete in under 0.02 seconds. Inputs up to ten million variants complete in under 10 minutes, demonstrating scalability from routine diagnostics to biobank scale analysis.

Table S1: **Standard evidence interpretation rules**. Each rule maps raw flags (TRUE, FALSE, NA) to binary evidence indicators based on whether the observed data contradict the underlying hypothesis. Rule names describe what a TRUE flag detects, and TRUE preferably corresponds to a contradiction or weakening signal. A value of 1 indicates that no contradiction is observed. Missing data (NA) never counts as evidence present. These rules assess verifiable evidence only; pathogenicity, penetrance, or counterfactual reasoning are handled upstream.

Rule name	**Flag**	**Meaning**	X_{ij}
parent gt unavailable	TRUE	Parental genotype data unavailable	0
parent gt unavailable	FALSE	Parental genotype data is available	1
parent gt unavailable	NA	Availability unknown	0
parent gt mechanistically inconsistent ar	TRUE	Parental genotype contradicts AR inheritance	0
parent gt mechanistically inconsistent ar	FALSE	No contradiction with AR inheritance observed	1
parent gt mechanistically inconsistent ar	NA	Parental genotype unavailable	0
compound het phase inconsistent	TRUE	Cis or phase contradicts compound heterozygosity	0
compound het phase inconsistent	FALSE	Trans or phase supports compound heterozygosity	1
compound het phase inconsistent	NA	Phase information unavailable	0
denovo mechanism excluded	TRUE	Parental genotypes exclude a de novo explanation	0
denovo mechanism excluded	FALSE	Parental genotypes do not exclude a de novo mechanism	1
denovo mechanism excluded	NA	Parental genotypes unavailable	0
unaffected relative homozygous alt	TRUE	Unaffected parent or close relative carries the same homozygous ALT genotype	0
unaffected relative homozygous alt	FALSE	No unaffected relative with the same genotype observed	1
unaffected relative homozygous alt	NA	Phenotype or genotype information unavailable	0

Table S2: **Required `QuantBayes` input format.** The first column must to contain a unique variant identifier, followed by the rule-based, m binary evidence indicators. This structure is required for both the command-line and R implementations. An example evidence feature is whether the proband genotype matches the required MOI.

variant_id	**rule1**	**rule2**	**rule3**	**...**	**rule** m
7-117559592-CTT-N	1	0	1	...	1
6-72183475-CG-N	0	1	1	...	0
1-14682421-G-A	1	1	0	...	1

Table S3: **QuantBayes output.** For each variant, k is the number of evidence rules satisfied, m is the total number of evaluated rules, θ is the posterior evidence sufficiency, and the 95% CrI and percentile describe its position within the genome wide distribution. The `variant id` column is shown here for illustration and is carried through from the input matrix; the raw text output omits this column and preserves the original row order. The global summary is reported in a separate `_global` file.

variant id	k	m	θ **mean**	θ **lower**	θ **upper**	**percentile**
7-117559592-CTT-N	18	24	0.7308	0.5487	0.8793	99.875
6-72183475-CG-N	18	24	0.7308	0.5487	0.8793	99.875
1-14682421-G-A	17	24	0.6923	0.5061	0.8505	99.375
...	...	...	...	...	...	...
Global	24	0.5192	0.5385	0.3846	0.6538	0.9500

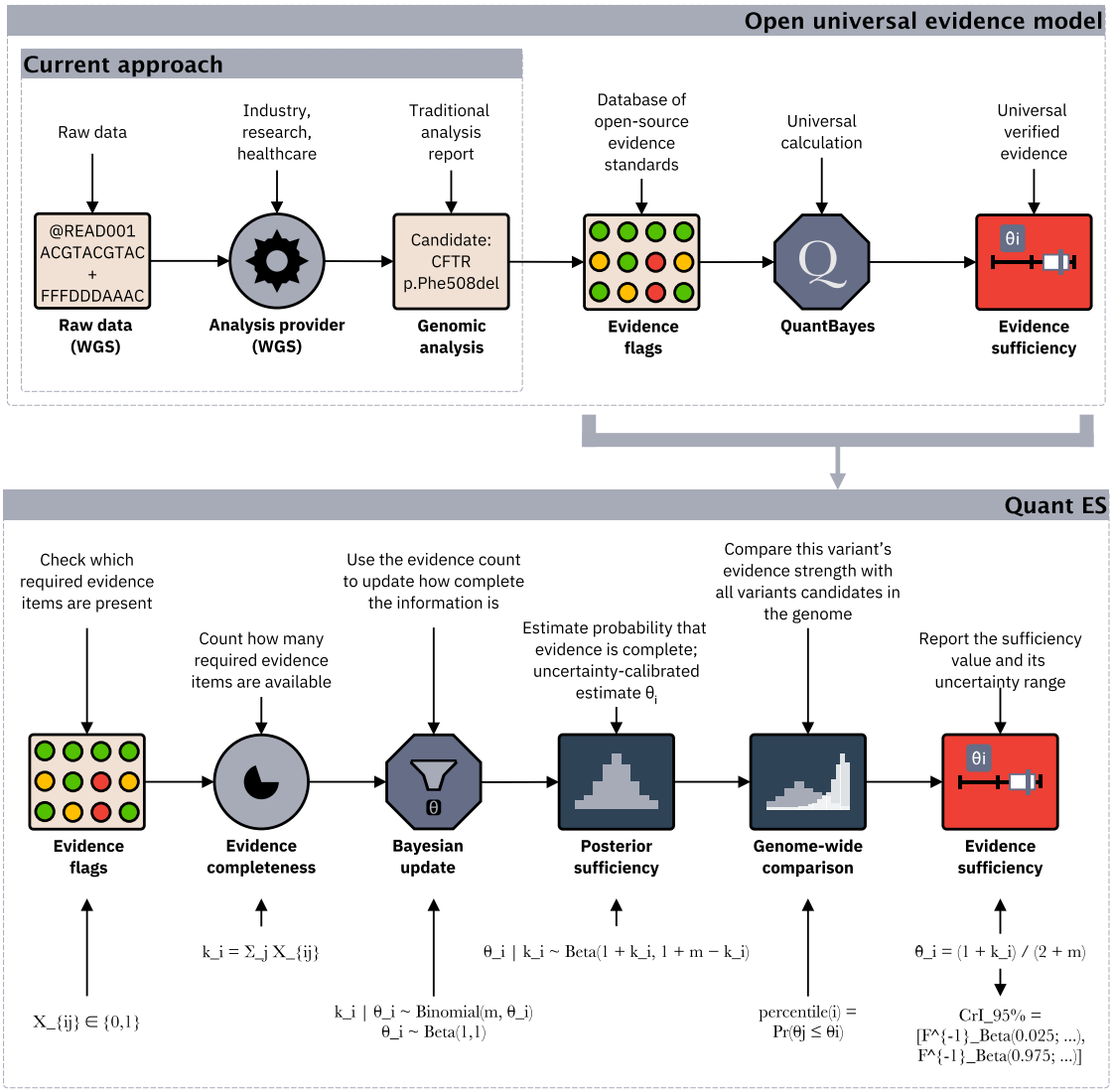


Figure S2: **Supplementary Quant ES calculation steps.** The schematic defines the symbols used in the Quant ES workflow. The input is a binary matrix X_{ij} with variants as rows and evidence features as columns. For variant i, the evidence count is $k_i = \sum_j X_{ij}$ and the number of evaluated features is m. The parameter θ_i denotes the evidence sufficiency for variant i. The figure shows the posterior $\theta_i \mid k_i$, its mean, its CrI, and the genome wide percentile $\Pr(\theta_j \leq \theta_i)$, which are the reported Quant ES outputs.

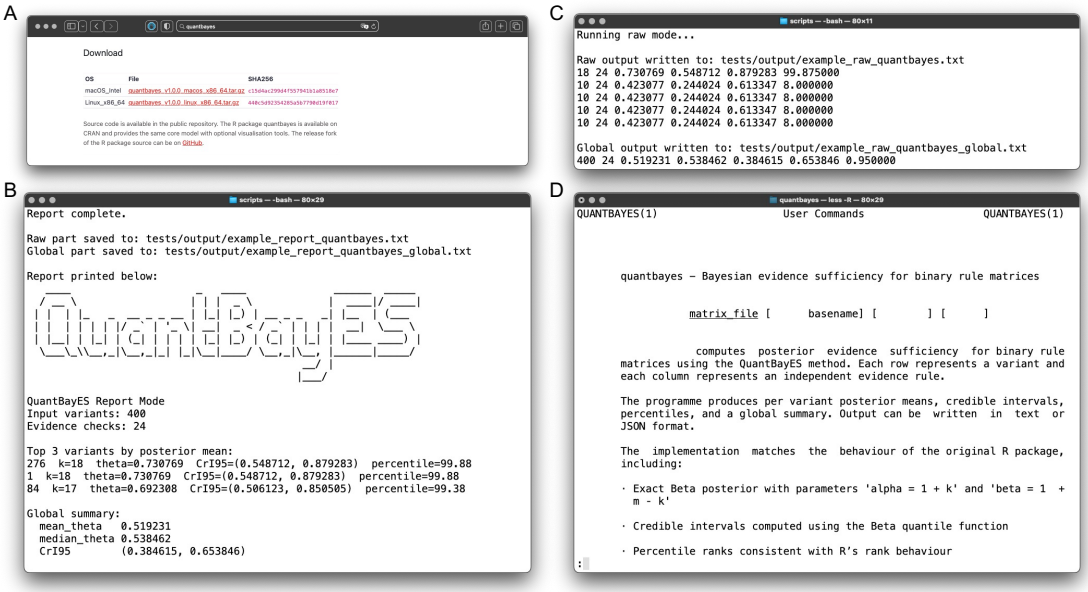


Figure S3: **QuantBayES example outputs.** (A) The cross-platform R package and executable binaries may downloaded from the public release archive for within research, clinical, or commercial environments. (B) Example console output produced using report mode, summarising top ranked variants and global mixture statistics. (C) Example raw output mode, showing per variant posterior means, CrIs, and percentiles. (D) Extract from the `QuantBayes` manual page illustrating usage examples.

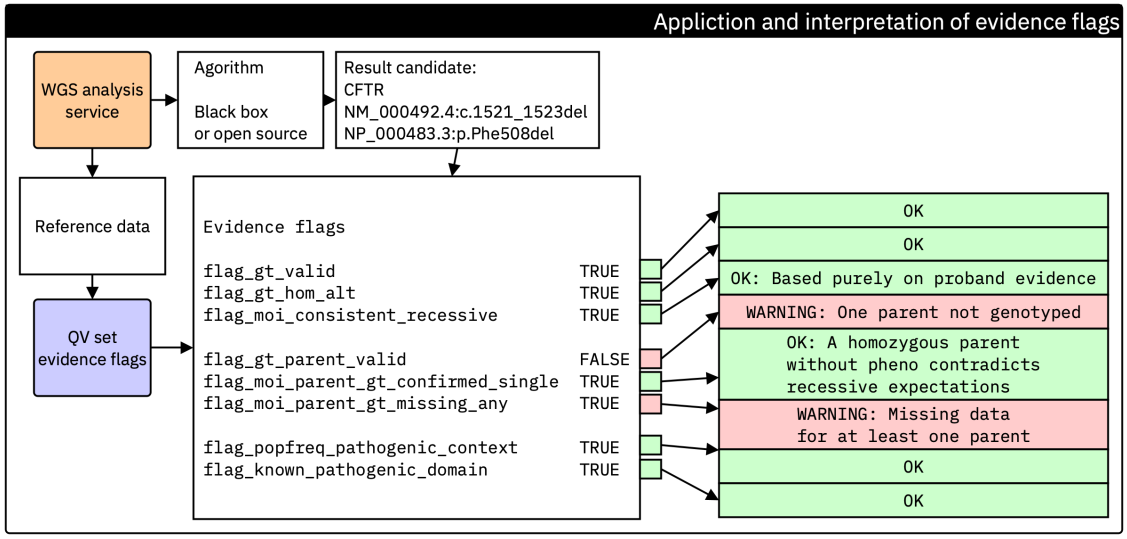


Figure S4: **Generation and interpretation of QV ES evidence flags.** After secondary analysis, each variant is evaluated against the registered evidence standard flag rule set. Each rule checks whether a required piece of biological or functional evidence is present. The results form a set of raw evidence flags (TRUE, FALSE, or NA).

Validation figures

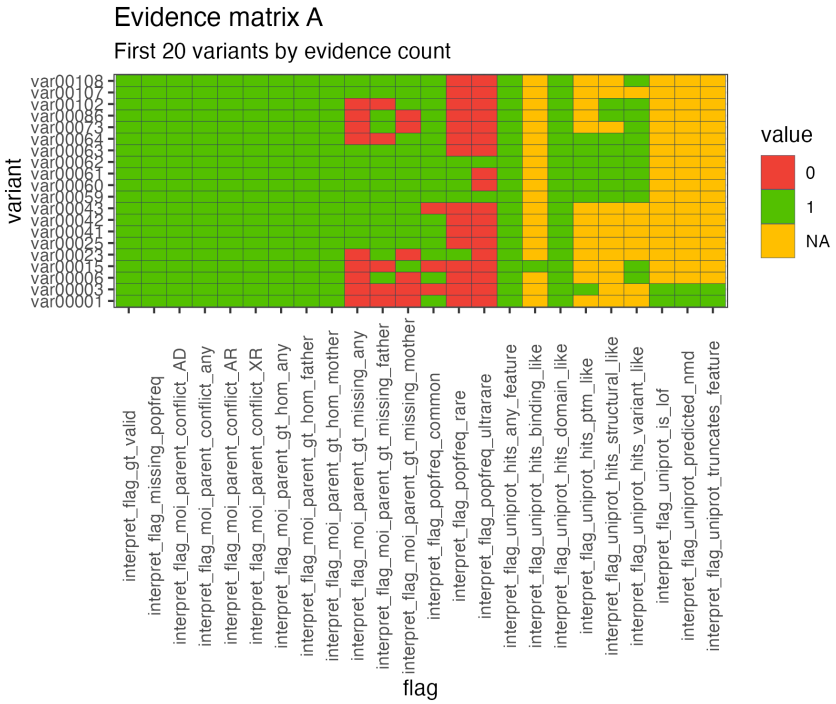


Figure S5: **Standard evidence interpretation matrix (GIAB WGS trio).** Each row corresponds to a variant and each column corresponds to a evidence standard rule. Cells encode the derived binary evidence outcome (1 = evidence present, 0 = evidence missing) after converting the Exomiser outputs using the flag set. The heterogeneity across rules illustrates the variability in available scientific evidence in a real whole genome analysis. We illustrate NA evidence in yellow indicating that this was specifically checked and no current evidence exists.

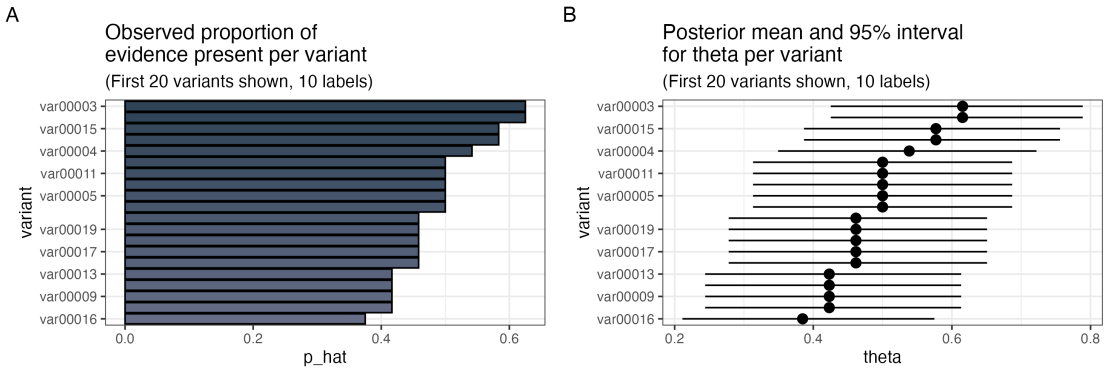


Figure S6: **Evidence sufficiency across variants (GIAB WGS trio). (A)** Observed proportion of satisfied evidence rules. Bars show $\hat{p}_i = k_i/m$ for the first 20 variants, illustrating the completeness of biological and functional evidence available for each locus. **(B)** Posterior means and 95% CrIs for θ_i. The Bayesian model converts evidence counts into uncertainty calibrated estimates of evidence sufficiency under real clinical analysis conditions.

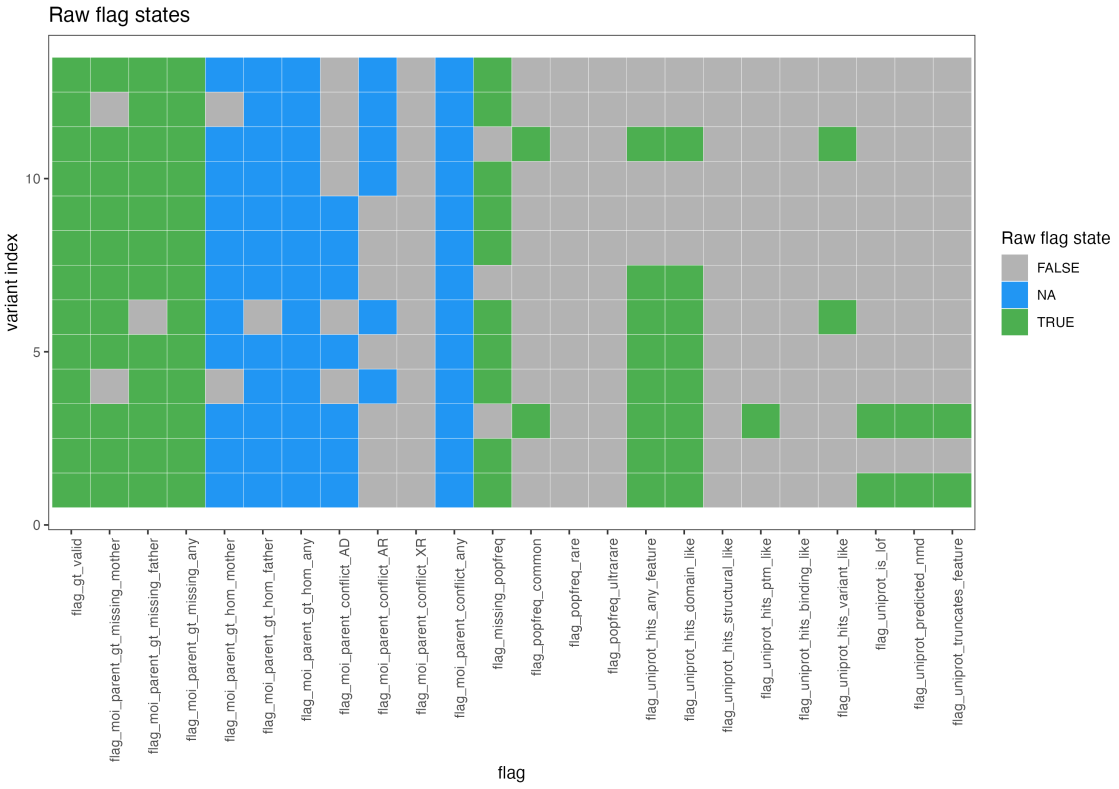


Figure S7: **Raw evidence flags from Exomiser outputs (GIAB WGS trio).** Rows represent variants and columns correspond to evidence standard rules. Cells show the direct flag outcomes returned by the evaluation: TRUE (evidence present), FALSE (evidence absent), and NA (evidence available within the dataset to confirm). These raw flags provide a standardised and provider independent summary of the evidence underlying each variant before statistical interpretation.

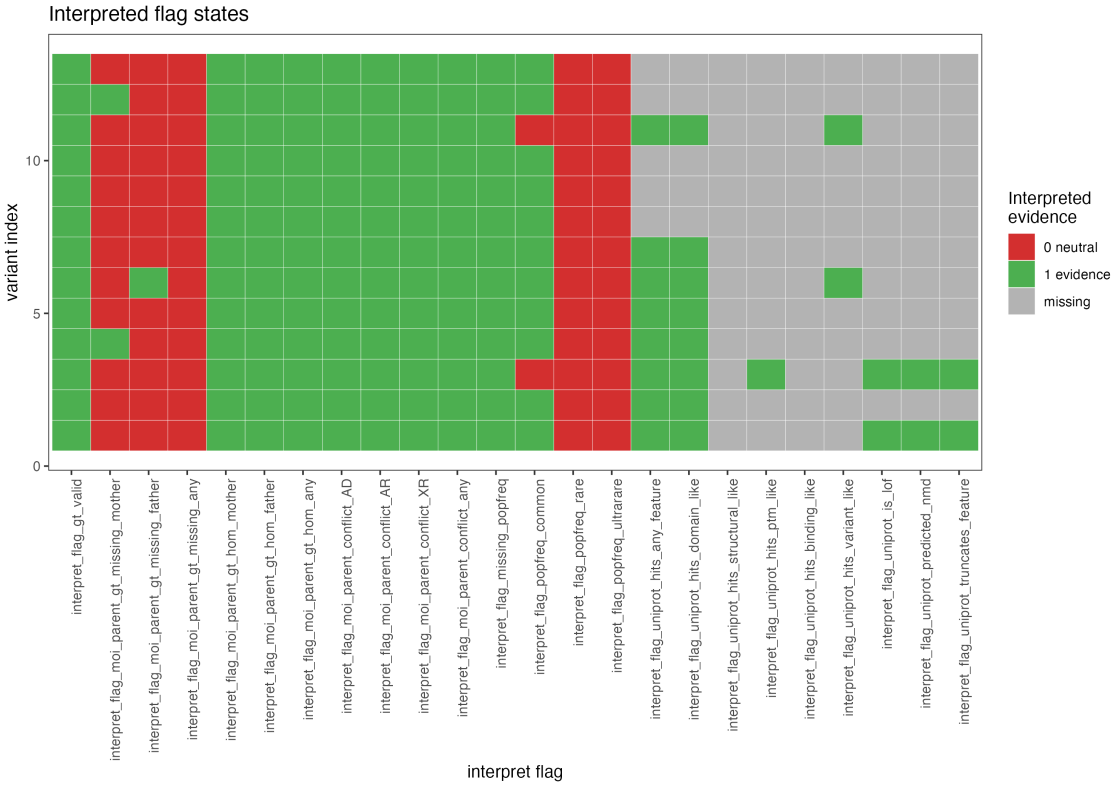


Figure S8: **Binary evidence interpretation matrix (GIAB WGS trio).** The registered interpretation rules convert the raw flags into a binary matrix suitable for Quant ES analysis. Cells indicate evidence present (1) or evidence missing (0), following the clinical interpretation logic of the evidence standard. The resulting matrix forms the input for Quant ES and enables consistent evidence comparison across all evaluated variants.

Statistical basis

Introduction

Clinical variant interpretation relies on multiple evidence domains whose availability varies substantially across the genome, directly affecting confidence in proposed causal variants. The qualifying variant evidence standard (8; 9) formalises this process as a finite set of verifiable evidence rules, each of which is either satisfied or not for a given variant. This structure naturally induces a binomial model of evidence availability.

The central analytical task is to convert a matrix of binary rule evaluations into a coherent probabilistic measure of how completely each variant is supported by available scientific knowledge. We address this using a Bayesian approach based on conjugate Beta–Binomial inference to estimate, for each variant i, a posterior evidence sufficiency probability θ_i, defined as the probability that a randomly selected evidence rule is satisfied. To provide contextual interpretation, each θ_i can also be evaluated relative to the empirical distribution of evidence sufficiency across all variants in the set, quantifying support alongside alternative plausible explanations (29).

The method is computationally efficient, scalable to millions of variants, and independent of external scoring systems or provider-specific heuristics. By separating evidence availability from analytic ranking, it yields a transparent, auditable, and reproducible statistical layer that complements existing clinical genetics practice.

Evidence structure and data model

For each variant i, we define a fixed set of m evidence rules $R_1, R_2, \ldots, R_m$, each representing a required and verifiable evidence domain. For variant i, rule j is encoded as

$$X_{ij} = \begin{cases} 1 & \text{if rule } j \text{ is satisfied,} \\ 0 & \text{otherwise.} \end{cases}$$

The resulting evidence vector is $\mathbf{X}_i = (X_{i1}, \ldots, X_{im})$, and the total number of satisfied rules is $k_i = \sum_{j=1}^{m} X_{ij}$. This observed evidence count reflects the completeness of available biological and functional knowledge for variant i. It does not represent pathogenicity, penetrance, or disease probability.

Variant-level evidence sufficiency model

We model the observed evidence count using a binomial likelihood $k_i \mid \theta_i \sim \text{Binomial}(m, \theta_i)$, where $\theta_i \in [0, 1]$ denotes the probability that a randomly selected evidence rule is satisfied for variant i. The parameter θ_i therefore quantifies evidence sufficiency rather than causal effect.

A Beta prior is assigned, $\theta_i \sim \text{Beta}(a, b)$, with default $a = b = 1$ (29). By conjugacy, the posterior distribution is $\theta_i \mid k_i, m \sim \text{Beta}(a + k_i,\ b + m - k_i)$, which under the uniform prior becomes $\theta_i \mid k_i, m \sim \text{Beta}(1 + k_i,\ 1 + m - k_i)$. The posterior mean is $\mathbb{E}[\theta_i \mid k_i] = \frac{1+k_i}{2+m}$, and uncertainty is summarised by the central 95% CrI

$$\left[\text{Beta}^{-1}(0.025; 1 + k_i,\ 1 + m - k_i),\ \text{Beta}^{-1}(0.975; 1 + k_i,\ 1 + m - k_i)\right].$$

This provides a calibrated measure of evidence completeness with explicit uncertainty.

Global evidence distribution

To contextualise variant-level evidence, we summarise the empirical distribution of θ_i across all evaluated variants. For each variant, posterior samples are drawn as

$$\theta_i^{(s)} \sim \text{Beta}(1 + k_i,\ 1 + m - k_i), \qquad s = 1, \ldots, S.$$

Pooling these samples yields a genome-wide distribution describing the overall availability of verifiable scientific evidence in the dataset. Global summaries, including the mean, median, and 95% CrI, characterise this background and provide context

for interpreting individual variants. This step is descriptive and does not alter the variant-level posterior inference.

Candidate comparison and relative evidence context

Variant interpretation is inherently comparative, but the inferential target of Quant ES remains local to each variant. After estimating the evidence sufficiency parameter θ_i independently for each variant, we contextualise candidates by comparing their θ_i values to those of all other evaluated variants. This is achieved through the empirical percentile; $\text{percentile}(i) = \Pr\big(\theta_j \leq \theta_i \mid j = 1, \ldots, N\big)$, where N denotes the total number of variants under the same evidence standard.

The percentile quantifies how well supported a candidate variant is relative to other observable alternatives. A high percentile indicates stronger evidential support within the evaluated set, while a low percentile indicates comparatively sparse support. This comparison is descriptive and does not constitute causal inference or prediction of pathogenicity. Although related to ideas of counterfactual comparison (30; 31), the percentile reflects relative evidence availability rather than the probability that a variant is the true causal explanation.

Behaviour under limiting evidence structures

Under the $\text{Beta}(1, 1)$ prior used in Quant ES, posterior certainty is intentionally bounded. Even when all m evidence rules are satisfied, the posterior

$$\theta_i \mid k_i = m \sim \text{Beta}(1 + m, 1)$$

has mean $(1 + m)/(2 + m) < 1$ and a credible interval that does not reach 1. Conversely, when no evidence is satisfied,

$$\theta_i \mid k_i = 0 \sim \text{Beta}(1, 1 + m)$$

remains strictly above 0. These properties prevent degenerate certainty in either direction.

The four simulated scenarios in **Table S4** illustrate extreme evidence configurations, with corresponding posteriors summarised in **Table S5**. When a variant has complete evidence, θ_i is high across scenarios, while its percentile depends entirely on the surrounding evidence landscape. When many variants share similar evidence patterns, percentiles cluster near the median; when competing variants lack evidence, the percentile approaches 100. Increasing the number of independent evidence rules sharpens the posterior for θ_i without altering this distinction. Together, these scenarios demonstrate that θ_i captures variant-level evidence completeness, whereas the percentile captures relative standing within the global evidence distribution.

In simple terms, if the overall level of evidence is low across the genome, this indicates that most variants lack interpretative information under the chosen evidence standard and that additional data or annotation may be needed. In this situation, a single variant with unusually high evidence support should be interpreted cautiously and may prompt a review of how evidence was gathered, particularly where manual input is involved, rather than being taken as proof of causality.

Table S4: **Input evidence patterns for the four limiting scenarios.** Each block shows the first three rows and first three evidence rules of the input matrix of 100 variants and 10 evidence features.

variant_id	**Scenario 1**	**Scenario 2**	**Scenario 3**	**Scenario 4**
variant 1	1 1 1	1 1 1	1 1 1	0 0 0
variant 2	1 1 1	0 0 0	1 1 1	0 0 0
variant 3	1 1 1	0 0 0	1 1 1	0 0 0

Table S5: Evidence sufficiency results for variant 1 in each of the four simulated scenarios. Calculation used a matrix of 100 variants and 10 evidence features.

Scenario	θ **CrI (95%)**	**Percentile**	**Global** θ **(95% CrI)**
1 variant 1	0.92 (0.72 – 1.00)	50.5	0.92 (0.92 – 0.92)
2 variant 1	0.92 (0.72 – 1.00)	100.0	0.09 (0.08 – 0.08)
3 variant 1	0.92 (0.72 – 1.00)	50.5	0.92 (0.92 – 0.92)
4 variant 1	0.08 (0.00 – 0.28)	50.5	0.08 (0.08 – 0.08)

Interpretation

A concise way to interpret Quant ES is to compare two evaluated variants. Suppose variant A satisfies most evidence rules, while variant B has many missing evidence indicators. Each variant is analysed independently from its own evidence count,

$$k_i = \sum_{j=1}^{m} X_{ij}, \qquad k_i \mid \theta_i \sim \text{Binomial}(m, \theta_i), \qquad \theta_i \mid k_i \sim \text{Beta}(1 + k_i,\ 1 + m - k_i).$$

The parameter θ_i summarises evidence completeness for variant i under the current evidence standard, and its CrI depends only on k_i and m. Consequently, sparse evidence for variant B does not widen the CrI for variant A. Quant ES does not estimate the probability that a variant is causal; it quantifies how complete the available evidence is. Alternative variants affect interpretation only through comparative context. Quant ES computes

$$\text{percentile}(A) = \Pr_{v_j \in \mathcal{V}}\big(\theta_j \leq \theta_A\big),$$

which places A within the distribution of evidence sufficiency across the evaluated set. A high percentile indicates that A is better supported by available evidence than most evaluated alternatives, but it does not exclude explanations outside the evaluated set.

Output

The per-variant output contains six columns. The values k and m record the number of supported evidence rules and the total number evaluated, respectively. The posterior mean of evidence sufficiency is reported as *theta_mean*, with uncertainty given by *theta_lower* and *theta_upper*. The *percentile* column reports the variant's rank within the genome-wide distribution of evidence sufficiency.

The global output summarises all evaluated variants. The fields *n_variants* and *m_rules* record matrix dimensions. The genome-wide evidence sufficiency distribution is summarised by *mean_theta*, *median_theta*, and the bounds *lower_theta* and *upper_theta* of the global credible interval, with *ci_level* indicating the interval probability.

Methodological advantages

The Bayesian evidence sufficiency framework has several practical strengths:

1. It separates biological knowledge availability from provider-specific analysis pipelines.
2. It quantifies uncertainty explicitly using closed-form posterior distributions.
3. It summarises variant support using principled evidence sufficiency and genome-wide context rather than arbitrary thresholds.
4. It is computationally lightweight and scalable to millions of variants.
5. It is fully auditable and reproducible through explicit evidence rules and outputs.
6. It does not encode pathogenicity directly, but instead quantifies the completeness of verifiable scientific information as a distinct interpretative layer.

Figure S9 provides a schematic overview of the Quant ES framework, showing how binary evidence rule outcomes are translated into posterior evidence sufficiency estimates and then contextualised within the genome-wide distribution.

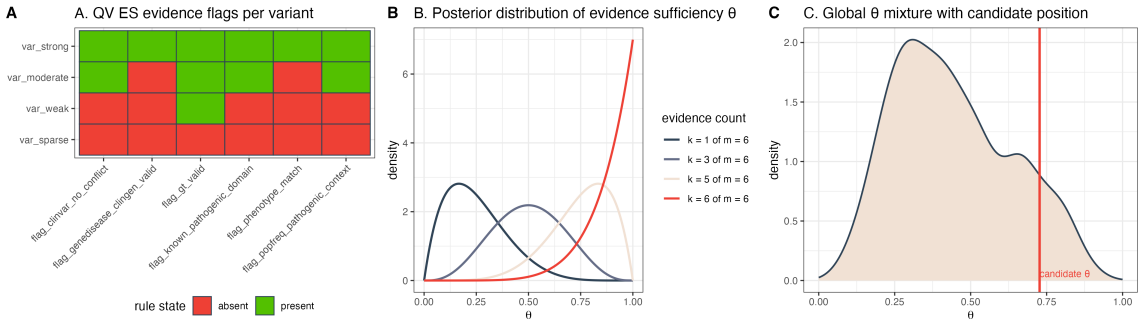


Figure S9: **Overview of the Quant ES Bayesian evidence model. (A)** an illustrative evidence matrix in which rows correspond to variants and columns to required evidence elements, with entries indicating whether each evidence requirement is satisfied. **(B)** Posterior distributions of the evidence sufficiency parameter θ for increasing numbers of satisfied evidence rules, illustrating how greater evidence completeness shifts posterior mass toward higher values. **(C)** The genome-wide distribution of θ together with the position of a selected candidate variant, situating its evidence completeness relative to all evaluated alternatives.

Summary

We presented a Bayesian framework for quantifying evidence rule satisfaction across genomic variants. Observed rule outcomes are modelled as binary observations arising from a latent evidence sufficiency parameter θ_i, which measures how complete the available biological and functional information is for variant i (17; 32). Each variant is further contextualised by its position within the genome-wide distribution of θ_i, yielding a percentile that reflects relative evidential support among evaluated alternatives. The approach is simple, scalable, and directly addresses a central constraint in clinical genomics: variation in the availability and completeness of verifiable evidence, rather than categorical assertions of pathogenicity.