Package: acmgscaler (via r-universe)

October 20, 2026

Title An R package for variant effect calibration to ACMG/AMP evidence levels

Version 1.0.0

Description Provides tools to calibrate functional variant scores to ACMG/AMP evidence strength levels for clinical variant interpretation.

License MIT + file LICENSE

Encoding UTF-8

Language en

Depends R (>= 3.5.0)

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.2

Repository https://badonyi.r-universe.dev

Date/Publication 2026-03-27 20:04:00 UTC

RemoteUrl https://github.com/badonyi/acmgscaler

RemoteRef HEAD

RemoteSha 641d35f4a3a7026e8a77c36976bc232d3f990eae

Contents

calibrate 2
variant_data 3

Index **5**

`calibrate` *Calibrate variant effect scores to ACMG/AMP evidence strength*

Description

The function calculates the positive likelihood ratio (LR, equivalent to the odds of pathogenicity) based on functional scores, e.g., from MAVEs or computational predictors, and their truthset labels. Score intervals for ACMG/AMP evidence levels are also computed. The input data requires at least one numeric column with the score of interest and another column, named `class`, with at least 10 pathogenic ('P') and 10 benign ('B') labels. Different or missing labels are allowed, but will be renamed to 'U'.

Usage

```
calibrate(df, value = NULL, prior = 0.1, group = NULL)
```

Arguments

`df`	A dataframe. Must have a `class` column with values 'P' (pathogenic) and 'B' (benign) labels, and a numeric column containing the variant effect scores. At least 10 occurrences of each class are required.
`value`	(optional) A character string indicating the name of the numeric column in `df` with the scores. If not provided, calibration will be run on all numeric columns.
`prior`	A scalar in the range 0-1 representing the prior probability of pathogenicity. Default 0.1.
`group`	(optional) A character string indicating the name of the column with the grouping variable. Default NULL.

Details

The function estimates the LR for each input score by resampling Gaussian kernel density estimates of the pathogenic and benign score distributions. Densities are mapped using linear interpolation and evaluated on a fixed-size common grid. To stabilise the LRs in regions where densities approach zero, a variance-based penalty is computed from log-LRs across 1,000 bootstrap replicates. This penalty is used to regularise the log-LR matrix. The log-LRs are monotonised in the principal direction of association with the input scores. Final estimates for each score include the point estimate and its 95% confidence interval. Score intervals for the different ACMG/AMP evidence levels are interpolated from the grid based upon the confidence bounds.

Value

A named list of dataframes. When grouping is not provided, the list has a length of two where 'likelihood_ratios' is the input dataframe with columns for LR and its confidence bounds (`column_name_lr`, `column_name_lr_lower` and `column_name_upper`). Assigned evidence classifications can be found in the `evidence` column. The second element in the list is named 'score_thresholds', which contain the lower and upper bounds of the score interval for ACMG/AMP evidence levels. When a grouping

variable is provided, the returned object is a nested list with a length equal to the unique group levels in the input data. Each of these elements contain the 'likelihood_ratios' and 'score_thresholds' dataframes.

References

Richards et al., 2015. Modeling the ACMG/AMP variant classification guidelines as a Bayesian classification framework. *Genetics in Medicine*. DOI: 10.1038/gim.2017.210

Tavtigian et al., 2018. 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*. DOI: 10.1038/gim.2015.30

Brnich et al., 2019. Recommendations for application of the functional evidence PS3/BS3 criterion using the ACMG/AMP sequence variant interpretation framework. *Genome Medicine*. DOI: 10.1186/s13073-019-0690-2

Pejaver et al., 2022. Calibration of computational tools for missense variant pathogenicity classification and ClinGen recommendations for PP3/BP4 criteria. *The American Journal of Human Genetics*. DOI: 10.1016/j.ajhg.2022.10.013

van Loggerenberg et al., 2023. Systematically testing human HMBS missense variants to reveal mechanism and pathogenic variation *The American Journal of Human Genetics*. DOI: 10.1016/j.ajhg.2023.08.012

Examples

```
# load example data provided with the package
library(acmgscaler)
data(variant_data, package = 'acmgscaler')

# compute likelihood ratios
calibrate(
  df = variant_data,
  # provide the column containing the variant scores (optional)
  value = 'score',
  # provide a grouping variable (optional)
  group = 'gene',
  # prior probability of pathogenicity (default 0.1)
  prior = 0.1
)
```

`variant_data` *Example variant data*

Description

Example data included in the package containing MAVE-derived functional scores and class labels for missense variants in BRCA1 and TP53. Functional scores are from high-throughput assays (Findlay et al., 2018; Giacomelli et al., 2018), and class labels are P/LP and B/LB from ClinVar.

Usage

```
variant_data
```

Format

A dataframe with 459 observations and 4 variables:

gene Gene symbol in which the variant occurs (e.g., `BRCA1`).

variant Missense variant, represented in single-letter amino acid notation (e.g., `L3F`).

class Binary class label indicating pathogenicity:

- `P`: Pathogenic
- `B`: Benign

score Functional assay score, typically representing the degree of functional disruption (numeric).

References

Giacomelli et al., 2018. Mutational processes shape the landscape of TP53 mutations in human cancer. *Nature Genetics*, 50(10), 1381–1387. DOI: 10.1038/s41588-018-0204-y

Findlay et al., 2018. Accurate classification of BRCA1 variants with saturation genome editing. *Nature*, 562(7726), 217–222. DOI: 10.1038/s41586-018-0461-z

Landrum et al., 2020. ClinVar: improvements to accessing data. *Nucleic Acids Research*, 48(D1), D835–D844. DOI: 10.1093/nar/gkz972

Index

∗ **datasets**
variant_data, 3

calibrate, 2

variant_data, 3