VASA USER MANUAL

Revision Sheet

Release No.	Date	Revision Description
Rev. 0	26/06/2017	Document creation (draft)
Rev. 01	17/08/2017	Updates to editing records and current worklist sections
Rev. 02	17/10/2017	Include multiple worklist sections
Rev. 03	24/11/2017	Changes to API section
v1.0	17/04/2025	

Table of Contents

1.0 INTRODUCTION

1.1 System Purpose

VASA (Variant Scoring Assistant) is a web-based tool that guides a user through the process of classifying variants against the American College of Medical Genetics and Genomics (ACMG) categories; recording evidence that supports decisions and calculating the pathogenicity score. The guidelines can be overridden, allowing scientists to exercise judgment where necessary.

VASA retains the calculations, variant annotations and associated evidence in a repository for reference in future submissions. It does not, however, store patient specific data, allowing sharing of evidence between users.

Variants may be scored multiple times and all scores and associated evidence are available to all users. Training VCF datasets can be uploaded allowing trainers to review how trainees scored variants, viewing the evidence recorded and which ACMG categories caused any variation.

1.2 Points of Contact

In the event of problems or questions, please contact the bioinformatics team on kch-tr.KCHBioinformatics@nhs.net.

1.3 Compatibility

VASA has been developed using Django as the web framework (version 3.2.25), and Bootstrap as the HTML & CSS framework (version v5.1.0), and as such should be compatible with most recent browsers.

At the time of writing, VASA has been testing with Google Chrome vw128, Mozilla Firefox v130 and Microsoft Edge 128.

2.0 GETTING STARTED

VASA is available at: https://vasa.kingspm.uk/

2.1 Logging On

Log in credentials can be obtained by emailing kch-tr.KCHBioinformatics@nhs.net. Please supply the name, email address and organisation of the user/s that require accounts.

2.2 Navigating the system

All functions are available from the navigation bar at the top every page (Figure 1).

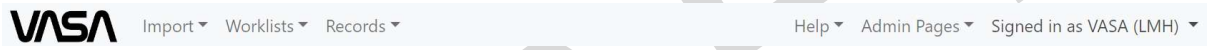


Figure 1. The navigation bar

Each function is covered in detail in sections 3 to 6; however, a brief overview is given below.

2.2.1 Import

This link gives a drop-down list to several methods for importing variants to the system to create a 'worklist' of variants to be scored. The options are:

- Import VCF: upload a locally stored VCF file to create a worklist.
- Import CSV: Import a locally stored comma-separated values (CSV) file.
- Import single variant: Manually enter a variant by genomic co-ordinates.
- Import training set: Create a worklist using one of the pre-loaded list of variants

If you are new to the system, choosing the 'example worklist' from the 'Import Training Set' will be the simplest way to get started and become familiar with the system.
Once a training set has been loaded, you will be redirected to the current worklist view described below.

2.2.2 Worklists

This displays either the current worklist or a list of previously imported worklists of variants being scored. The table displays the progress of scoring for each variant and links to initiate scoring.

2.2.3 Records

This links to existing scores contained within the VASA database.

2.2.4 Help

This contains links to the application details, the user manual and contact details for technical support.

2.2.5 Login/Signed in as….

This is the link for logging in, or if you are already logged in, the link for logging out or changing your password.

3.0 IMPORTING A VARIANT LIST

A variant, or variant list, can be imported using several methods. This section describes how to manually upload variants to create a worklist. Uploading variants via the API is covered in section 6.

With the exception of the 'training set' import, you will be presented with a form as shown in Figure 2. The common fields are described below.

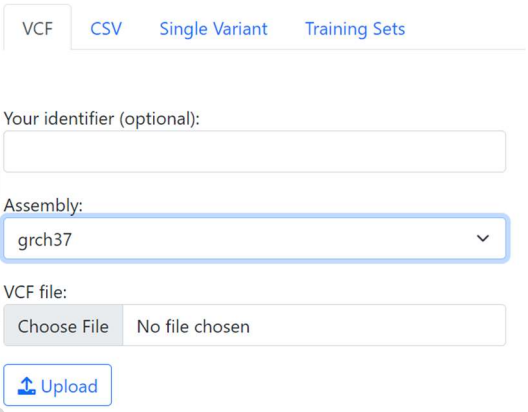


Figure 2. VCF file import dialogue box

Your identifier (optional): This field will be stored with the final variant score and it is intended to provide a link between the user's sample and the final score (i.e. a laboratory identifier) when searching the VASA database. It should not be used for patient identifiable information.
If this option is left blank, VASA will use the file name as the identifier if a VCF or CSV is uploaded, or simply the genomic coordinates in the case of an individual variant.

Assembly: This defaults to GRCh37 and can be changed to GRCh38.

Choose a file: Enter the path of the local copy of the VCF file you wish to import.

Once a variant or list of variants have been successfully imported, you will be redirected to the worklist page (see section 4).

3.1 Import VCF

VASA uses VCFPy 0.13.8 (https://vcfpy.readthedocs.io/en/stable/) to parse the VCF file and so supports VCF version 4.3.

VCF Annotation

VASA has been designed to accept annotated VCF files using VEP annotation (http://grch37.ensembl.org/Homo_sapiens/Tools/VEP), and tested upto Ensembl version 90 (including the new gnomAD annotation). While this is optional, the advantage of using annotated VCF files is that the data will be available during the scoring process to users and will be stored with the final scores in the database. This can assist the user in the scoring process and provide an audit trail of the evidence that was available at the time of scoring, without the need to manually input data.

There is no defined set of expected VEP annotation fields; VASA will display and store whatever annotation fields are provided.

** If more than one annotation set is provided for a variant, an entry for each will be displayed in the worklist

3.2 Import CSV

Like the VCF file, VASA can be used to import comma-separated values (CSV) files of variants. As with the VCF files, VASA will parse the associated variant annotation and store this with the variant record. The required fields are: CHROM, POS, REF, ALT and transcript.

3.3 Import a Single Variant

You can add a single variant for scoring for *ad hoc* tasks. The form the variant's genomic co-ordinates in the format:

[chromosome]:[position]:[reference]:[alternative] e.g: 1:123456:A:T

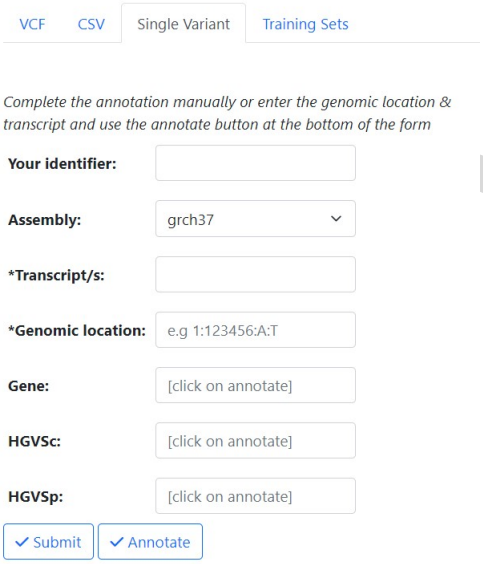


Figure 3. Single variant import dialogue box.

VASA will make a request to Ensembl to confirm the validity of the co-ordinates and the reference allele. The system will work with GRCh37 and GRCh38 human genome assemblies.

A transcript ID must also be provided (eg 'NM_001754.5'). To annotate the variant with HGVS nomenclature and gene information from VariantValidator, click 'Annotate' before submitting the variant. There is no validation of the transcript field, so if a variant sits outside a transcript, you can enter '?' for the transcript.

Possible errors:	
Please check the format of the genomic co-ordinates	The variant has not in the correct format described above
Unable to parse HGVS notation	The variant is correctly formatted but is outside of the expected range (i.e. an invalid chromosome number or co-ordinate)
Reference allele extracted from xxx does not match reference allele given by HGVS notation	The co-ordinates are valid but the reference allele is not the one expected. If this occurs, in the first instance check the use of the correct assembly

3.4 Import a Training Set

VASA has several built-in sets of variants, both for use as a training tool for the application itself, or for training on scoring variants with the ACMG guidelines.
If you are new to the system, the 'example worklist' will be a good place to get started and become familiar with the system.
Most training sets will come with information at the top of the worklist page as additional instructions, such as clinical background in the case of hypothetical training cases.

Previous scores will not be available for training variants and the scores you submit will be hidden. You can view any training variants you have scored by navigating to Records > My training set submissions.

To add a new training set, please contact kch-tr.KCHBioinformatics@nhs.net.

3.5 The Worklists

VASA will store worklists to allow users to swap between them and work on more than one dataset at a time (Figure 4). Any worklist that is deleted will not impact on variants already submitted to the database, but you will lose any partially scored variant data.

Your Saved Worklists

	Session ID	Date Created	Date Modified	Your identifier
	10008	13 Aug 2024, 11:08 a.m.	13 Aug 2024, 11:08 a.m.	Neuro Somatic Training Set
	9732	2 Aug 2024, 9:55 a.m.	2 Aug 2024, 9:55 a.m.	sqvd
	9704	1 Aug 2024, 12:37 p.m.	1 Aug 2024, 12:37 p.m.	sqvd
	9675	31 Jul 2024, 4:19 p.m.	31 Jul 2024, 4:19 p.m.	sqvd
	9674	31 Jul 2024, 4:18 p.m.	31 Jul 2024, 4:18 p.m.	sqvd
	8757	4 Jul 2024, 2:15 a.m.	4 Jul 2024, 2:15 a.m.	6245817_4:186603546 C>T
	8697	2 Jul 2024, 4:33 p.m.	2 Jul 2024, 4:33 p.m.	AML training set
	8696	2 Jul 2024, 4:32 p.m.	2 Jul 2024, 4:32 p.m.	LMH Training set [retired]

Figure 4. The user's worklists stored in VASA

4.0 SCORING VARIANTS

4.1 Current Worklist

Once a list has been imported, you will be redirected to the worklist view. This lists all the imported variants, any associated annotation and links to score the variant and show if this variant has been seen before by VASA. This is shown in

Figure 5.

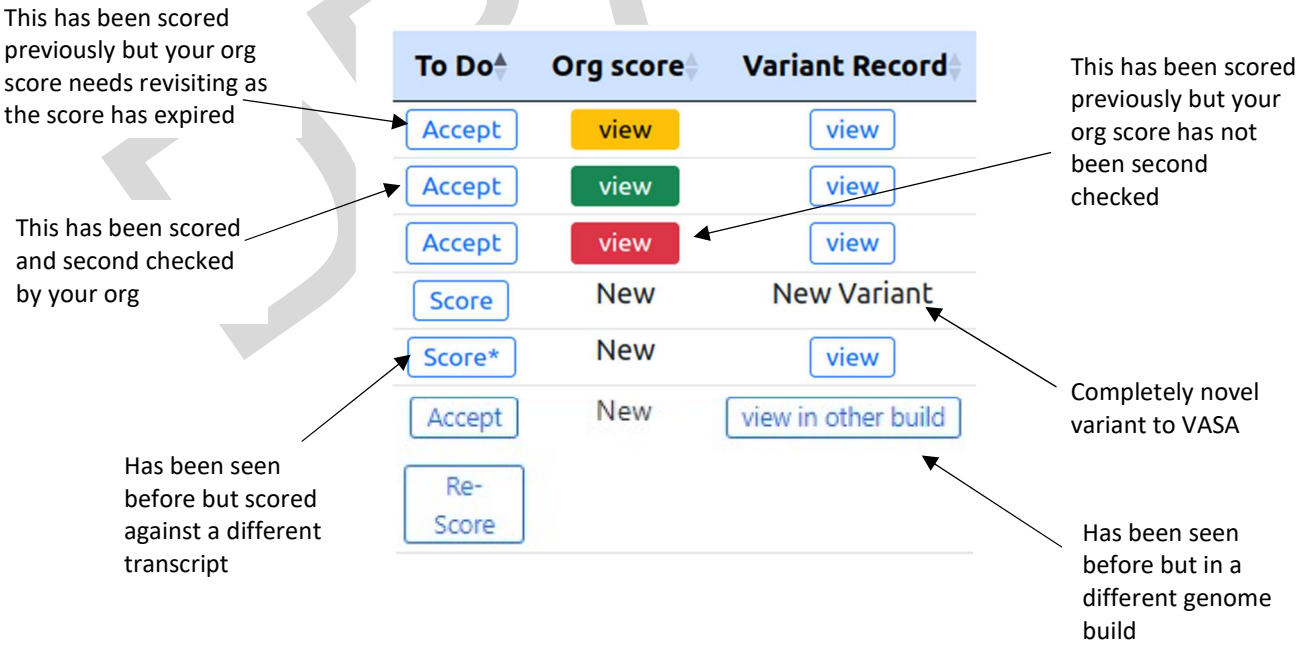


Figure 5. The current worklist view. The variant data and annotation has been cropped, and only the first three columns have been shown here.

The To Do column

The worklist will keep track of the variants you score. The 'Score' button allows you to start work on a variant and 'In progress' shows which variants have already been

started. ‘Done’ indicates a variant has been scored and submitted to the database in this session.
If a variant has been previously scored by your organisation for the same transcript version, there will be the option to ‘accept score’. Choosing this option will not change anything in the database, but will simply allow you to track your progress within the worklist.

VASA will allow you to work on multiple variants and retain progress on each. Progress will be kept even if you log out and/or close the browser.

The ORG Score column

This will show if a variant has been scored previously by your organisation, along with a link to the variant record.
The button is also colour-coded:

- green means score has been second checked,
- red means no second check
- amber means needs review (the second check was > 365 days ago).

The Variant Record column

If the same variant has been scored (against a different transcript or by a different organisation), this will link to the variant stats page with an overview of all record/s.

4.2 Scoring a Variant

From the worklist view, click on the white ‘Score’ button to start scoring your chosen variant (or ‘In Progress’ to continue with a variant has been partially scored).

4.2.1 Category Dashboard

This will take you to the ‘category dashboard’ showing the 28 ACMG categories and other summary information, as shown in Figure 6.

Figure 6. The category dashboard

All the categories will start as ‘Pending’ (coloured orange) and you can work though answering as many or few categories as you consider relevant.

The categories are grouped by type of evidence (population database, functional studies, family data, etc.).

The overview panel to the right of the screen shows the current score based on the categories that have been completed. As per the ACMG guidelines, a variant defaults to ‘VUS’ in the absence of evidence of pathogenicity or benignity.

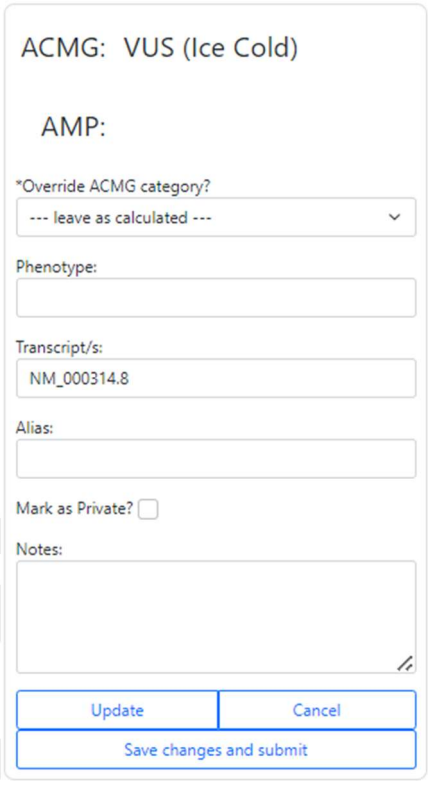


Figure 7. The Overview panel.

The Phenotype field

This field is compulsory, and is a free-text field that will populate with HPO terms as suggestions. Multiple HPO terms can be entered separated by a comma.

Alias

This is an optional field for noting any alternative, historic names for the variant.

Transcript

This is a required field for the transcript the variant is being scored against. This field is validated against the existing variant database to check there is not a previous score for this variant and transcript version within your organisation. Multiple transcripts or transcript versions can be entered, separated by a comma.

Override final category?

It is possible for the scorer to override the final variant score if required. If this is done, the score will be displayed with an information button as shown below:

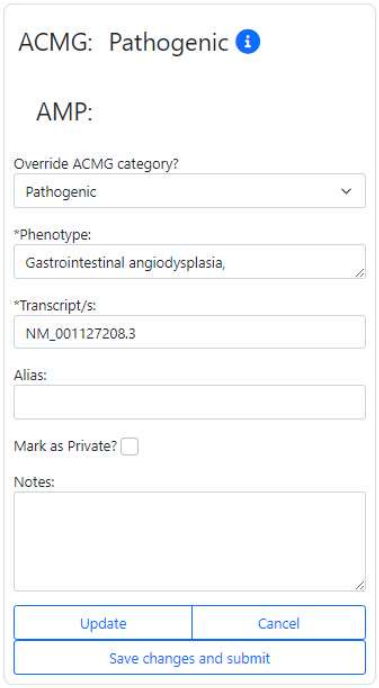


Figure 8. Overview panel with amended score.

Notes can also be added that will be stored with the variant scoring.

4.2.2 Category Questions

Clicking on the any category button will take you to the page for storing information for each category. This is shown below in for the first category, BA1.

The top panel gives the full text for the category from the ACMG guidelines, with two panels below that. The lower left panel is for recording information relating to the category and the lower right panel will display variant annotation (if provided) that is pertinent to the category in question.

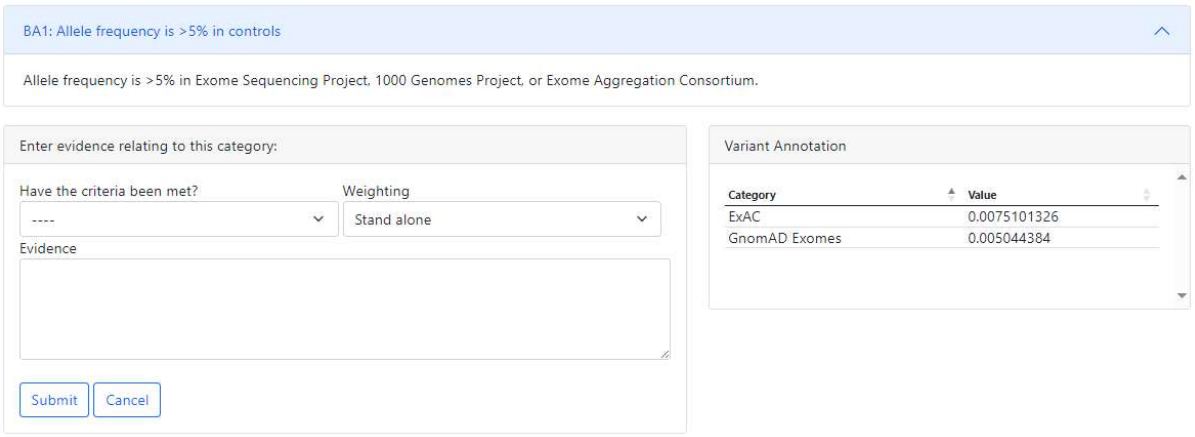


Figure 9. The category question page

Have the criteria been met?

This gives three options; ‘yes’, ‘no’ and ‘Not known’. Checking ‘yes’ will mark this criterion as met, with regards to calculating the variant’s pathogenicity. The other two options will **not** change the scoring of the variant, but will allow the user to (optionally) record that the category has been considered and either not met or there is insufficient evidence. This may be useful in more complex cases where the scorer wished to record that they have considered the criterion but believe it has not been met.

Weighting

This dropdown list allows the option to change the default weighting for the category, as there may be some occasions where this is required. The new weighting for the category will be applied when the pathogenicity score is calculated.

Evidence

This is an optional free-text field for recording comments relevant to this category.

Once the category information has been recorded, clicking submit will return you to the dashboard and the category will be marked as changed, as shown below. You can return to any category to amend or update information until the final submission.

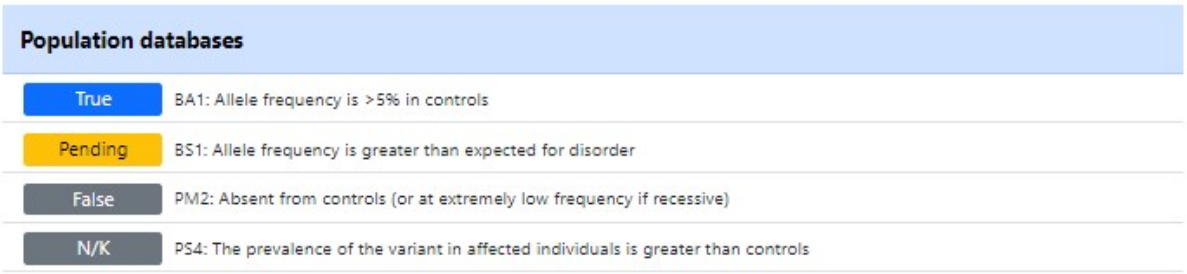


Figure 10. The display will change once a question has been completed.

4.2.3 AMP Scoring

This tab allows the option to also the variant using the AMP scoring criteria. The options available are:

- Not Required
- Tier I
- Tier II
- Tier III
- Tier IV

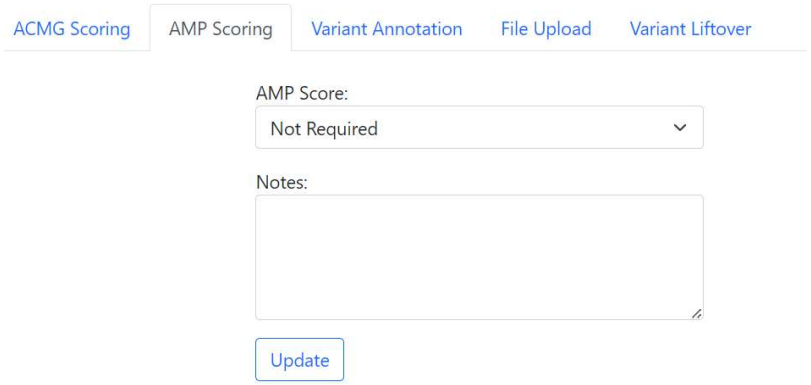


Figure 11. AMP Scoring panel

4.2.4 Variant Annotation

This tab allows you to view all the variant annotation that was imported with the variant. This may be just the core variant details (chromosome, genomic position, etc.) or more depending on the options at import.

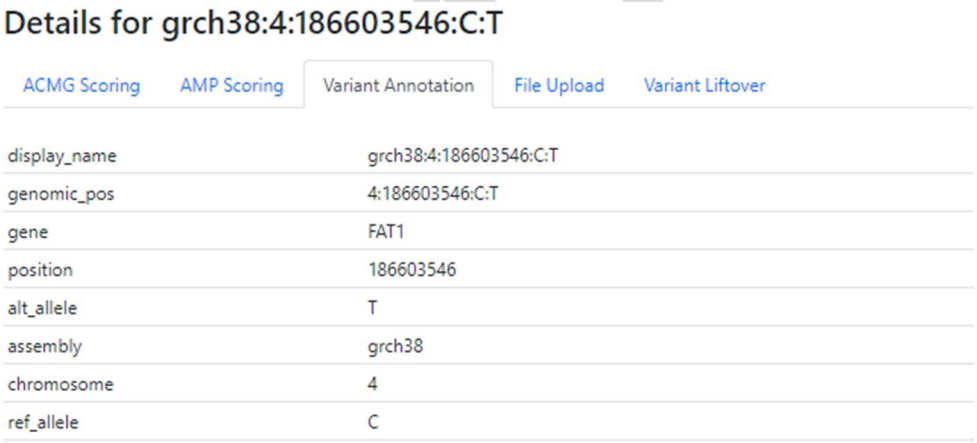


Figure 12. Variant Annotation

4.2.5 File Upload

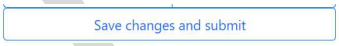
This allows you to upload any files that are associated with your variant scoring.

4.2.6 Variant Liftover

This tab provides you the genomic coordinates for the variant in both GRCh37 and GRCh38. By clicking the 'Review liftover' button you can check VariantValidator annotation for this variant and confirm the liftover between genome builds is correct.

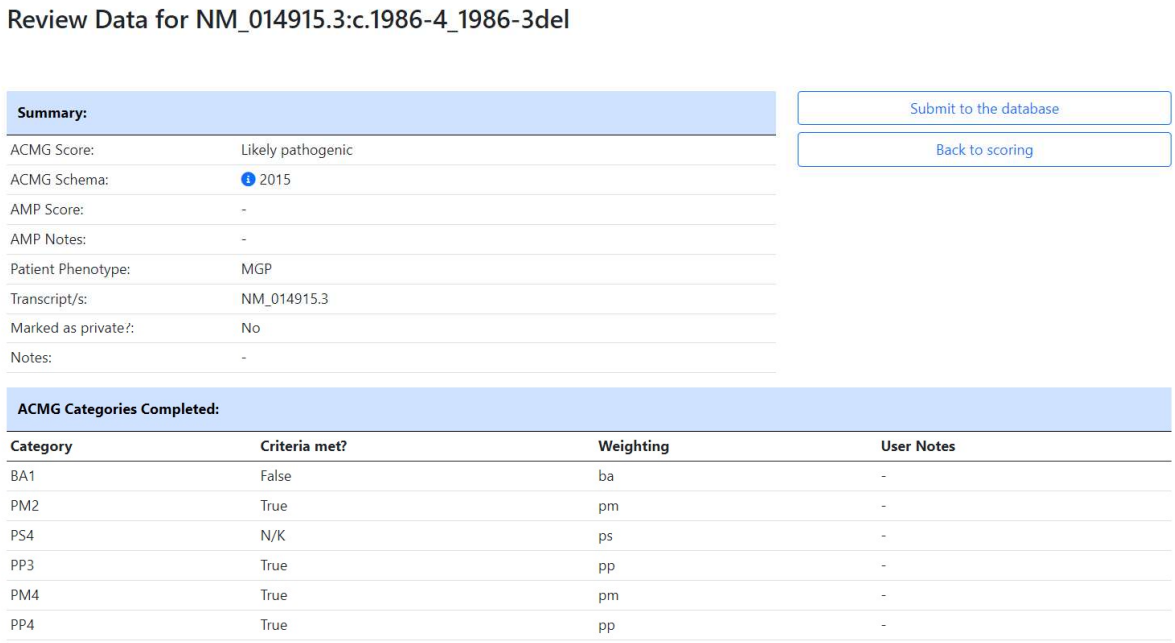
Figure 13. Liftover of variant genomic coordinates.

4.2.7 Reviewing and Submitting the Score

Once you are satisfied the scoring is complete, click the 

button.

This will take you to a summary page with an overview of the AMCG categories completed during scoring and any notes entered against them:



Summary:	
ACMG Score:	Likely pathogenic
ACMG Schema:	2015
AMP Score:	-
AMP Notes:	-
Patient Phenotype:	MGP
Transcript/s:	NM_014915.3
Marked as private?:	No
Notes:	-

ACMG Categories Completed:

Category	Criteria met?	Weighting	User Notes
BA1	False	ba	-
PM2	True	pm	-
PS4	N/K	ps	-
PP3	True	pp	-
PM4	True	pm	-
PP4	True	pp	-

Figure 14. The review page.

Nothing will be committed to the database until the ‘Submit to the database’ button is clicked. VASA will check if the variant is already present in the database, and if so, if will associate the new score with that variant. If the variant is new to the database, it will submit both the variant and score data.

Once submitted, you will be returned to the ‘current worklist’ page.

4.2.5 Amending an existing record

When viewing an existing record during the scoring process, if the score was submitted by your organisation, there will be 'Edit Record', 'Edit notes' and 'Update annotation' buttons at the top of the page.

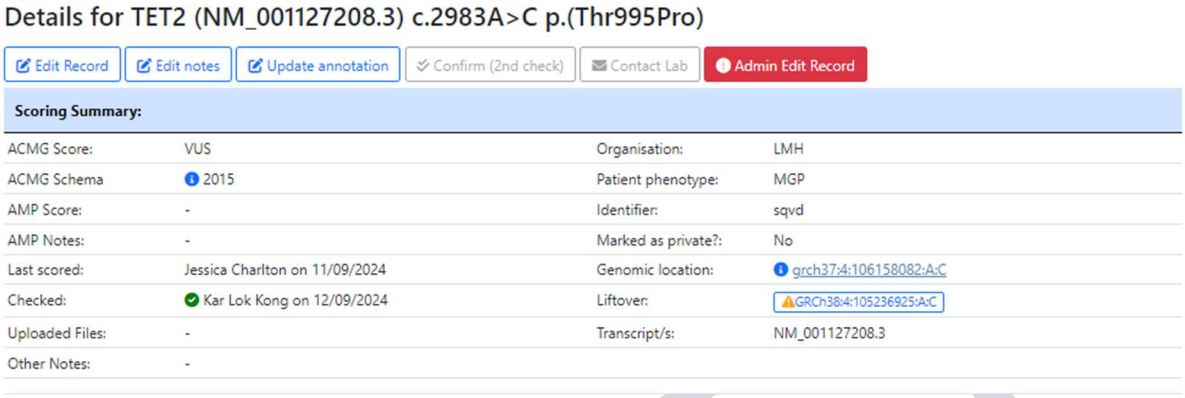


Figure 15. Editing a record

The 'Edit Record' button will open the record in the scoring dashboard again and allow you to amend the record and resubmit. Please note,

- If you choose to edit a record that you did not originally submit, the record will be amended to record you as the new scorer of the record.
- If the record had been second checked, this status will be cleared and will now need a second check from another user in your organisation. Thus, any changes to a record will now need a second person to confirm those changes.

The 'Edit notes' button will open a pop-up to allow you to add notes to the record or phenotype information without changing the last scorer or second checked status of the record:

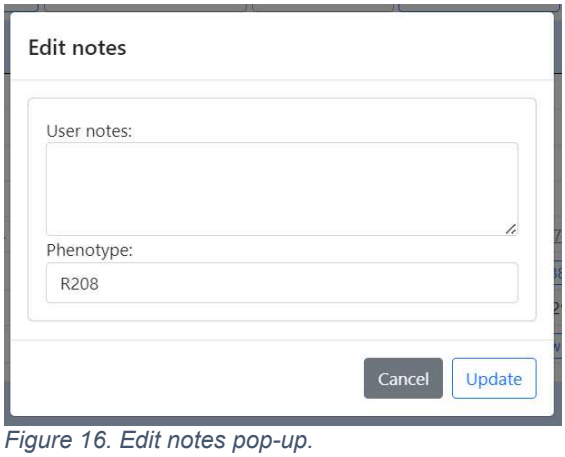


Figure 16. Edit notes pop-up.

The 'Update Annotation' button does not change the annotation information stored with the variant record at the time of scoring or the second check status but will allow

you to update the title/display name at the top of the page. This button queries VariantValidator using the transcript associated with the variant score and returns a new display name in the format: gene (transcript) HGVSc HGVSp.
After checking the result is correct, click ‘Update’ to save the changes.

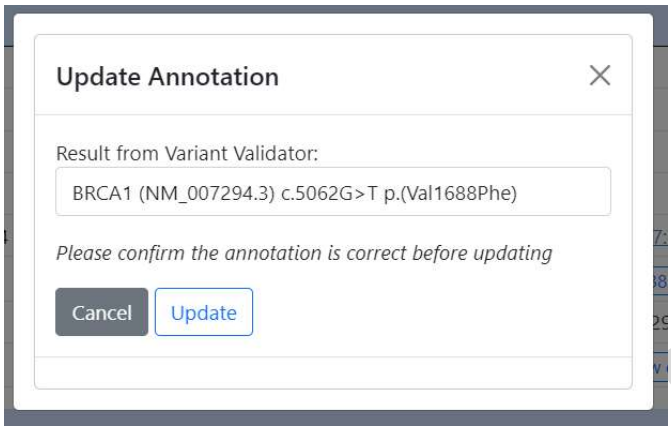


Figure 17. Update Annotation pop-up.

If the record has not been second checked (and you were not the original scorer), there will be a second button for you to confirm the data and you will be marked down as the second checker for the record. A pop-up will be displayed for you to confirm you are happy to confirm the second check.

A history of previous scoring can been found at the bottom of the variant record page:

Variant Scoring history

ACMG Score	ACMG Schema	AMP Score	Scored by	Date scored	Second check	Date checked
Benign (amended from VUS) current	2015	-	nahid.kamal@nhs.net	22/11/2022	stevenbest@nhs.net	22/11/2022
Likely benign (amended from VUS)	2015	-	l.ficinski@nhs.net	26/10/2021	stevenbest@nhs.net	17/11/2021
Likely benign (amended from VUS)	2015	-	l.ficinski@nhs.net	26/10/2021	-	-
Likely benign (amended from VUS)	2015	-	l.ficinski@nhs.net	26/10/2021	-	-
Benign (amended from VUS)	2015	-	l.ficinski@nhs.net	30/08/2021	-	-
Benign (amended from VUS)	2015	-	karlok.kong@nhs.net	21/08/2020	stevenbest@nhs.net	21/08/2020

Figure 18. Variant scoring history.

5.0 VIEWING THE RECORDS

There are three basic views for querying the data stored in the records: ‘Search all records’, ‘Export records’ and ‘My training set submissions’.
Although activity such as scoring a variant is logged against a user in the database, VASA will only display the user’s organisation (not the named individual) to database viewers from outside that organisation. For your own organisation, you will be able to see the name and contact email for the users’ who scored and second checked a record.

5.1 All Scored Variants View

A screenshot of the variants table is shown below. It gives the details for each variant score and annotation information such as gene and HGVSp (if provided), as well as phenotype, the ACMG score and when the score was submitted and second checked.

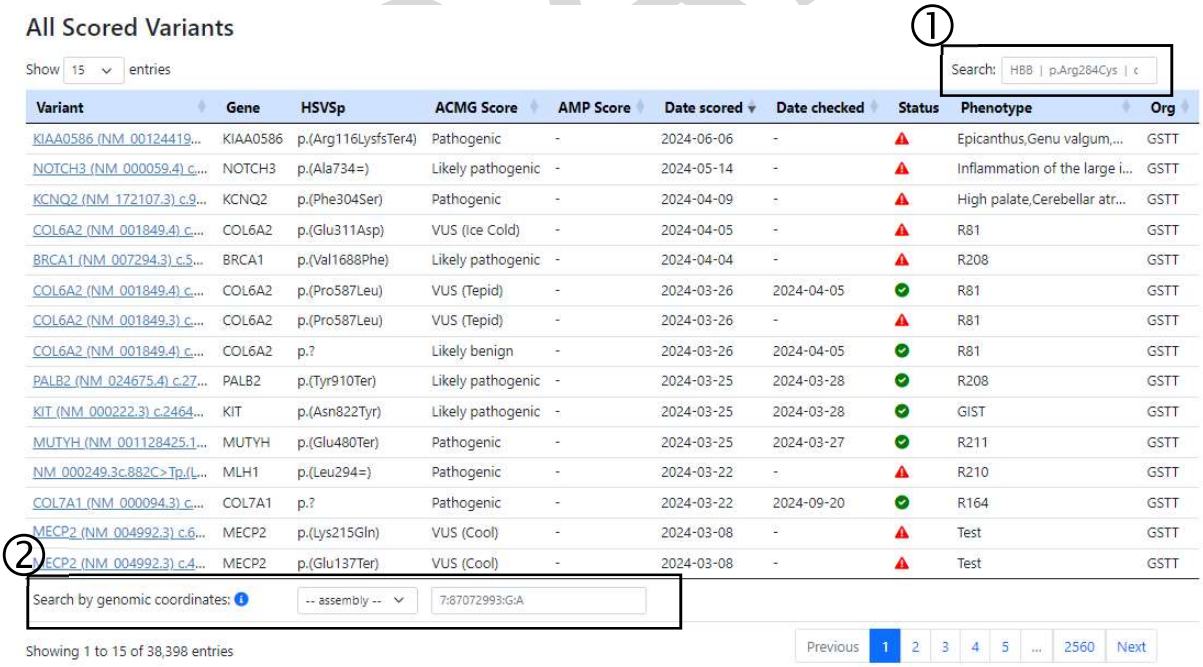


Figure 19. All scored variants view.

There are two search functions in the all scored view:
① Free text search – you can search for variant scores by gene, HGVSc, HGVSp, and genomic coodinates.
② Genomic coordinate search – genome assembly must be provided and a liftover is performed to check for variant scores in either GRCh37 or GRCh38.

There is a link to view more details for the variant score record. This is described in section 5.2

5.2 Scoring Summary

Selecting one of the variants from the 'All Scored Variants' table will give more detailed data on that variant score record. This page gives the details for each submitted score, including the audit information such as the scorer and date and time plus all the submitted categories answers and user submitted evidence. If it was provided, it will also store the variant annotation available at the time. This should prove useful as some annotation may change over time, but it will be possible to see a snapshot of the evidence provided at the time of scoring.

Scoring Summary:

ACMG Score:	VUS	Organisation:	LMH
ACMG Schema	2015	Patient phenotype:	Myeloid Gene Panel
AMP Score:	-	Identifier:	-
AMP Notes:	-	Marked as private?:	No
Last scored:	Jessica Charlton on 11/09/2024	Genomic location:	grch37:X:39914681:A:T
Checked:	Kar Lok Kong on 12/09/2024	Liftover:	GRCh38:X:40055428:A:T
Uploaded Files:	-	Transcript/s:	NM_001123385.2
Other Notes:	manual import of legacy MGP data. Remains VUS, JC 11/09/24.		

ACMG Categories Completed:

Category	Criteria met?	Weighting	User Notes
PM1	False	-	Not located in a functional domain or hotspot.
PM2	True	pm	exomes: not found (cov: 68.0) genomes: not found (cov: 23.4)
PP3	False	-	REVEL: Uncertain, 0.578 CADD: Pathogenic Supporting, 26.7
PP5	N/K	-	Nothing in COSMIC or ClinVar and no relevant literature identified.

Figure 20. Score details

To view all scores submitted for a variant by all organisations, you should click the genomic location (highlighted in the above screen shot by red box).

5.3 Variant Details

This page provides detailed variant information, including when it was first submitted to the database and a link to all scores submitted against it by all organisations.

5.3.1 Summary

This shows the genomic location of the variant, how many times the variant has been scored, and the average and most frequent score.

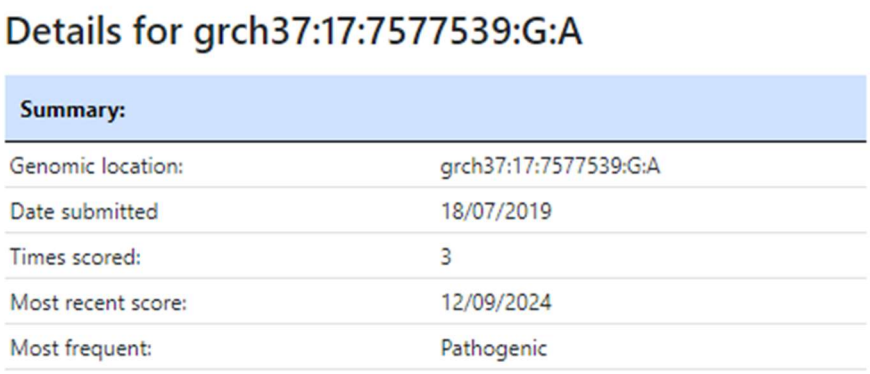


Figure 21. Variant Details

5.3.2 Individual scores

This provides a list of the all previous scores and a link to view each score (further details on individual score views are in section 5.4). Click the icon to be taken to the variant scoring summary for that score record.

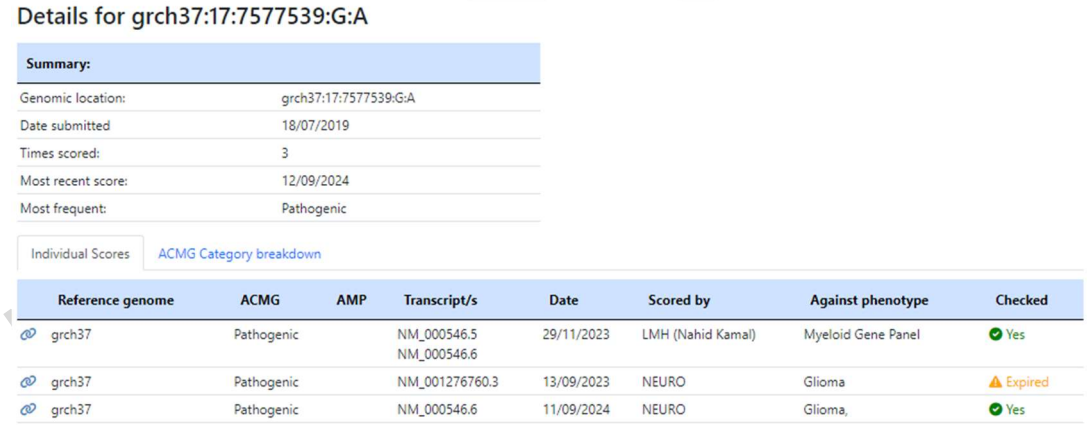


Figure 22. A list of previous scores per variant.

5.3.3 ACMG Category Breakdown

This shows a summary of which ACMG categories have been marked as met across all the submitted scores for that variant. This will give a useful insight into the origins of variation in the final scoring of variants:

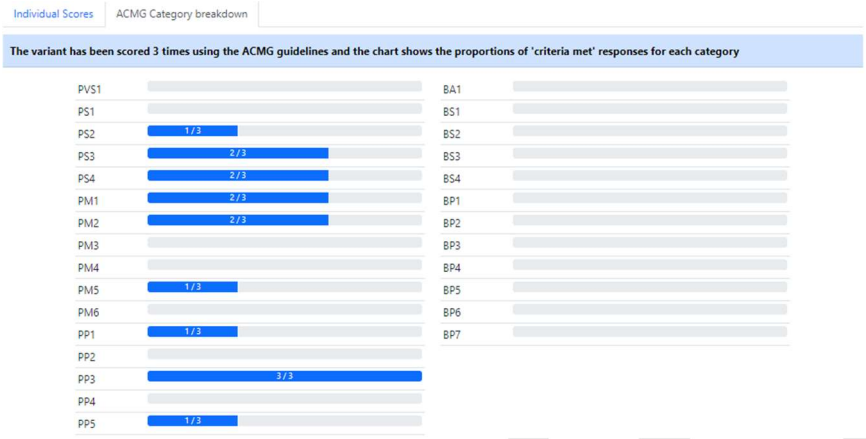


Figure 23. Category summary breakdown.

6.0 API REFERENCE

A basic API is available to provide a simple, automated method integrating VASA into your existing workflows. API endpoints are available for creating worklists and retrieving variant scores. The VASA API schema is available in the VASA swagger pages here: https://vasa.kingspm.uk/rest/v3/swagger/. A brief summary of the endpoints is also available below.

6.1 Token authentication

For the following sections, 6.3 & 6.4, simple user token authorisation has been implemented. The request will need to contain the authorisation in the header:

```
Authorization: Token <user's token>
```

A user's token can be found in two ways. Either from the GUI or programmatically. From the website, click on the 'signed in as…' in the menu bar, then 'view user profile'. To retrieve the API token programmatically. Using a command line http client, such as httpie:

```
http POST li446-39.members.linode.com/api-token-auth/ username='user'
password='password'
```

VASA will return the user's token in a JSON response:

```
{
    "token": "<user's token>"
}
```

6.2 Submitting a worklist

You can submit a complete worklist to VASA, including any annotation. Below is a simple python script demonstrating the process:

```
import requests

url = https://vasa.kingspm.uk/rest/v3/post_worklist/'
auth_token = 'Token <your token>
headers = {'Authorization': auth_token, 'content-type': 'application/json'}

worklist = '''
{
  "run_name": "worklist1",
  "assembly": "grch37",
  "annotation_type": "vep",
  "variants": [
    {
      "chromosome": "X",
      "position": "107930929",
      "ref_allele": "G",
      "alt_allele": "C",
      "transcript": "NM_000495.4",
      "display_name": "COL4A5(NM_000495.4)c.4510+5G>C",
      "gene": "COL4A5",
      "consequence": "splice_region_variant&intron_variant"
    }
  ]
}
'''

response = requests.post(url, data=worklist, headers=headers)
print(response.json())
```

The example above only includes the core information but as many annotation fields can be included as desired. The 'annotation_type' is used by VASA to determine the types of annotation.
VASA returns the id of the worklist in a simple JSON format, thus allowing the creation of a URL to redirect the user within the requesting application:

```
{'url': 'http://[hostname]/imports/worklist/xx'}
```

6.3 Querying the database

The Django REST Framework has been used to create the API for querying data. Please refer to section 6.2 for obtaining your API token.

Individual variant records can be retrieved with:

```
https://vasa.kingspm.uk/rest/v3/variant_detail/[genomic_position]/?format=json
```

Individual scores can be retrieved with:

```
https://vasa.kingspm.uk/rest/v3/score_detail/[score index]/?format=json
```

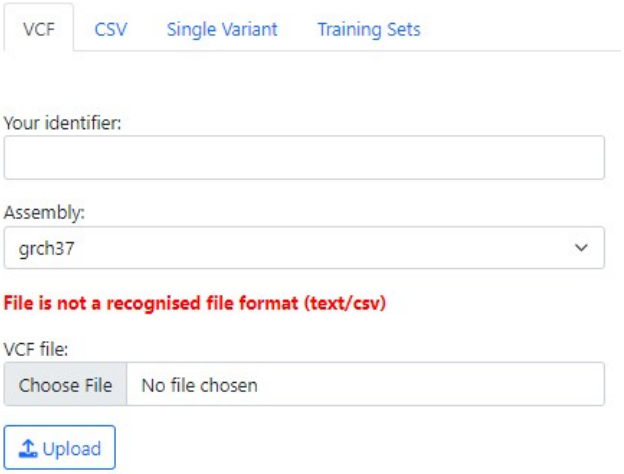
The format for records can be seen on the following page.

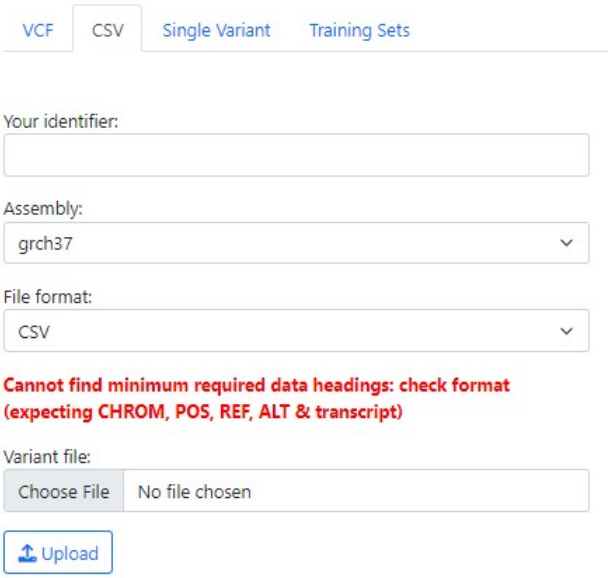
```
"pk"           : int,                                  # VASA index number
"chromosome"   : string,                               # Following columns refer to variant submission in db
"position"     : string,
"ref_allele"   : string,
"alt_allele"   : string,
"genomic_pos"  : string,                               # in '1:g.123456A>T' format
"assembly"     : string,
"date_created" : string,
"submitted_by" : string,
"organisation" : string,
"scores": [                                            # 'scores' is a list: multiple scores per variant
    {
        "pk"             : int,                        # VASA index number
        "final_score"    : string,                     # Final score (the 'auto_score' may have been overridden)
        "auto_score"     : string,                     # Automatic scoring based on ACMG criteria
        "phenotype"      : string,
        "user_notes"     : string,
        "scored_by"      : string,
        "organisation"   : string,
        "date_submitted" : string,
        "date_modified"  : string,
        "scored_by_acmg" : boolean,                    # If false the following scoring_data field is ignored
        "scoring_data"   : {                           # scoring_data holds ACMG category data. A child object for each category
            string : {                                 # The category name, e.g. 'PVS1', 'BA1'
                "user_text"         : string,
                "weighting"         : string,          # e.g. pvs, ps, pp, ba, (this may have been changed from default)
                "category"          : string,
                "criteria_met"      : True|False|N/K
                "initial_weighting" : string           # e.g. pvs, ps, pp, ba, (the default, not necessarily the value used)
            }
        },
        "variant_annotation" : {                       # Variant annotation: any number of key/value pairs. This is per score, not variant
            Key: Value,
        },
    },
]
```

7.0 TROUBLESHOOTING

Importing Variants

If VASA is unable to parse the CSV or VCF file of variants you will receive the following error page:





If this happens, please confirm the file is in the correct format. If the problem persists, please see section 1.2 for support contact details.

8.0 ACRONYMS

CSV	Comma Separated Values
VEP	Ensembl's Variant Effect Predictor
VCF	Variant Call Format
ACMG	American College of Medical Genetics and Genomics
AMP	Association for Molecular Pathology
HGVS	Human Genome Variation Society