Final version after processing the results of discussion in the working group						**Colour Legend**			
This will become a collection of all tables; after reviewing per item						**Row colours**		**mandatory/recommended/optior**	
						Confirmed elements		mandatory and mandatory if applicable	
Review the sheet Sequence									
Please leave comments including your name and date						Newly added elements - for review		recommended	
If done with review (even if no remarks/comments) add your name below under Reviewers of this version								optional	
Version history; Items added per version								conditional	
Items	Date version release								
Sequence	2025-01-30								
Authors of latest version:									
Name	Date version release								
Ana Konrad	2025-01-30								
Hannah Neikes	2025-01-30								
Jeroen Belien	2025-01-30								
Joeri van der Velde	2025-01-30								
Abhishek Nayak	2025-01-30								
Aedin Culhane	2025-01-30								
Reviewers of latest version									
Name	Date								
David Salgado	2025-02-05								
Mikael Kronborg	2025-02-10								
Priit Kleemann	2025-02-10	UTARTU expets are also reviewing and validating this table, can give their feedback at task meeting.							
Richard Hagan	2025-02-10								
Evita Lindholm	2025-02-10								
Hannele Laivuori	2025-02-11								
Milan Ojsteršek	2025-02-11								
Michela Riba	2025-02-11								
Gabriele Bucci	2025-02-11								
Edel Cahalin	2025-02-10								

Suggested_Union_Domain/class/table	Suggested Union Item	Suggested_Union_Proposed_Definition	Suggested_Union_Proposed_Values	Suggested_Union_Proposed_cardinality	Suggested_union_mandatory/recommended/optional	If conditional, terms of condition stated here	Part of sunflower:	Link to terminology/ontology that defines item	Link to terminology/ontology that defines value(s) /valuelist(s)	Reasoning/explanation/evidence of/for suggestion		Concept: Exact match with	Concept: close match with	Example data (completely synthetic)		Link to terminology/ontology that defines the class	Class: Exact match with	Class: close match with
Sequence	**Target**	Identification of the sequenced target	Whole genome sequencing Whole exome sequencing Multi-gene panel sequencing array OTH	1..1	mandatory	A lab protocol is always applicable, for conventional sequencing but also for microarrays, optical genome mapping, karyotyping etc.	Core	https://bioportal.bioontology.org/ontologies/EFO/?p=classes&conceptid=http%3A%2F%2Fpurl.obolibrary.org%2Fobo%2FOBI_0000272	Whole genome sequencing Whole exome sequencing Multi-gene panel sequencing array OTH	A lab protocol is always applicable, for conventional sequencing but also for microarrays, optical genome mapping, karyotyping etc.						http://www.sequenceontology.org/browser/current_release/term/SO:0000110		
Sequence	**Sequencing_Date**	Defines the date of sequencing.	Date, ISO 8601 format, YYYY-MM-DD	0..1	optional				https://www.iso.org/iso-8601-date-and-time-format.html									
Sequence	**ISO 15189 accredited**	Indication whether the laboratory is accredited according to ISO 15189 (clinical)	boolean	0..1	recommended		Core		NA									
Sequence	**ISO 17025 accredited**	Indication whether the laboratory is accredited according to ISO 17025 (research)	boolean	0..1	recommended		Core		NA									
Sequence	**WIP** **Wet lab protocol**		String or https://www.protocols.io URL		optional		Core		String or https://www.protocols.io URL	**Note: This item is still work in progress!**								
Sequence	**Participated in proficiency testing**	Indication whether the laboratory had participated in any proficiency testing, such as interlaboratory comparison	boolean	0..1	conditional	If not accredited for either ISO standard.	Core		NA									
Sequence	**IVDR_passed**	Indicate whether the methodology (including chemistry and sequencing standards) used for sequencing follows the In vitro diagnostic medical devices (IVDR) regulation passed by the EU in April 2017.	boolean	0..1	optional				NA	http://data.europa.eu/eli/reg/2017/746/oj								
Sequence	**Sequencing_Platform**	The used sequencing platform (i.e. brand, name of a company that produces sequencer equipment).	FAIR Genomes or EFO list	1..n	mandatory	A sequencing platform, machine or device is always applicable, for conventional sequencing but also for microarrays, optical genome mapping, karyotyping etc.	Core	https://bioportal.bioontology.org/ontologies/NCIT/?p=classes&conceptid=http%3A%2F%2Fncicb.nci.nih.gov%2Fxml%2Fowl%2FEVS%2FThesaurus.owl%23C172274	FAIR Genomes or EFO list	A sequencing platform, machine or device is always applicable, for conventional sequencing but also for microarrays, optical genome mapping, karyotyping etc.								
Sequence	**Average depth of coverage**	Mean coverage for whole genome sequencing, or mean target coverage for whole exome and targeted sequencing (eg 60x, average number of times each target base has been 'read' by sequencer).	Integer	1..1	mandatory		Core	https://ncicb.nci.nih.gov/ncimbrowser/ConceptReport.jsp?dictionary=NCI%20Metathesaurus&code=CL555947	NA	read depth, sequence depth or depth of coverage, refers to the number of times a specific base (nucleotide) in the DNA is read during the sequencing process. In other words, it's the average number of times a given position in the genome is sequenced. A higher sequencing depth provides more confidence in the accuracy of the base calls at that position and helps to reduce sequencing errors and noise.								
Sequence	**Breadth of coverage**	Breadth of coverage (or evenness) is the proportion or percentage of a reads that has been sequenced at a the provided average depth of coverage. (E.g. if for Average depth of coverage, the value is 60, and the Evenness is 50% at 60x, the value here will be 50)	Integer	0..1	recommended		Core	Not found yet	NA	Coverage is the proportion or percentage of a genome that has been sequenced at a certain depth. It gives an idea of how much of the entire genome has been effectively read and is usually expressed as a multiple of the genome's size.								
Sequence	**Additional NGS quality control metrics**	Statement of any additional NGS quality control metrics	String	0..1	optional				NA									
Sequence	**Initial_input_file_format**	Identification of the genomic file format of the initial input file (eg. fastq, bam, cram)	**EDAM's filetypes and format**	1..1	mandatory	A resulting digital file is always applicable, for conventional sequencing but also for microarrays, optical genome mapping, karyotyping etc.	Core	https://bioportal.bioontology.org/ontologies/NCIT/?p=classes&conceptid=http%3A%2F%2Fncicb.nci.nih.gov%2Fxml%2Fowl%2FEVS%2FThesaurus.owl%23C171252	**EDAM's filetypes and format**	A resulting digital file is always applicable, for conventional sequencing but also for microarrays, optical genome mapping, karyotyping etc.								
Sequence	**Final_output_file_format**	Identification of the genomic file format of the final output file (eg. vcf, gvcf)	**EDAM's filetypes and format**	1..1	mandatory	A resulting digital file is always applicable, for conventional sequencing but also for microarrays, optical genome mapping, karyotyping etc.	Core	https://bioportal.bioontology.org/ontologies/NCIT/?p=classes&conceptid=http%3A%2F%2Fncicb.nci.nih.gov%2Fxml%2Fowl%2FEVS%2FThesaurus.owl%23C171252	**EDAM's filetypes and format**	A resulting digital file is always applicable, for conventional sequencing but also for microarrays, optical genome mapping, karyotyping etc.								
Sequence	**Final_output_file_format_version**	Identification of the version of genomic file format of the final output file (eg. vcf, gvcf)	**String**	0..1	recommended				NA									
Sequence	**Alignment_software**	Identification of the software used for alignment	**Digital Resource**	0..n	recommended	If aligned file format is provided; alignment software has to be stated	Core	https://bioportal.bioontology.org/ontologies/NCIT/?p=classes&conceptid=http%3A%2F%2Fncicb.nci.nih.gov%2Fxml%2Fowl%2FEVS%2FThesaurus.owl%23C175896	NA	Conditional on file format: does not apply to FASTQ files or other results that are independent of alignment								
Sequence	**Alignment_Genome**	The specific build of the human genome used as reference for sequence alignment and variant calling.	**Digital Resource**	1..n	mandatory	If aligned file format is provided; alignment genome has to be stated	Core	https://bioportal.bioontology.org/ontologies/NCIT/?p=classes&conceptid=http%3A%2F%2Fncicb.nci.nih.gov%2Fxml%2Fowl%2FEVS%2FThesaurus.owl%23C164815&jump_to_nav=true	NA	Conditional on file format: only applies to files dependent on alignment such as BAM and VCF								
Sequence	**Specific_Settings_Alignment_Genome**	Any specific settings regarding alternative contigs or decoys.	String	0..n	optional		Core		NA									
Sequence	**Targeted_Gene**	In case of targeted sequencing, specify which gene is being targeted. This item points to another class: Target_Gene	**Target_Gene**	0..n	optional			https://bioportal.bioontology.org/ontologies/OBI/?p=classes&conceptid=http%3A%2F%2Fpurl.obolibrary.org%2Fobo%2FOBI_0001962	NA									
Sequence	**Target_Other**	Any other targeted genomic region	**Children of** http://www.sequenceontology.org/browser/current_release/term/SO:0000001 **NULL flavors**	0..n	optional				**Children of** http://www.sequenceontology.org/browser/current_release/term/SO:0000001 NULL flavors									
Sequence	**Panel_of_Normals_Included**	Indicate whether a panel of normals is included during variant calling	**boolean**	0..1	recommended				NA									
Sequence	**Panel_of_Normals_Description**	Free text description of panel of normals, if applicable.	**string**	0..1	conditional	If panel of normals == T			NA									
Sequence	**Variant**	A detected and reported variant	**Variant**	0..n	optional				NA									
Sequence	**Variant_calling**	Identification of the software used for variant calling	**Digital Resource**	0..n	conditional	If panel of normals == T			NA									
Sequence	**Variant_calling_date**	Defines the date of variant calling.	Date, ISO 8601 format, YYYY-MM-DD	0..1	optional				https://www.iso.org/iso-8601-date-and-time-format.html									
Sequence	**Variant_Annotation**	Identification of the software used for variant annotation	**Digital Resource**	0..n	recommended	If variant calling is performed, software used for variant annotation stated here			NA	Conditional on file format: only applies to files with genomic variation such as VCF and gVCF.								
Sequence	**Variant_Annotation_database**	Database and version used for variant annotation	**Digital Resource**	0..n	recommended	If variant calling is performed, database for variant annotation is to be stated		https://bioportal.bioontology.org/ontologies/FDAM/?p=classes&conceptid=http%3A%2F%2Fedamontology.org%2Foperation_3225	NA	Conditional on file format: only applies to files with genomic variation such as VCF and gVCF.								
Digital Resource	**Name**	The name of the tool/software/database used.	String	1..1	mandatory				NA							http://www.ontologydesignpatterns.org/ont/dul/DUL.owl#DigitalResource		

Suggested_Union_Domain/class/table	Suggested Union Item	Suggested_Union_Proposed_Definition	Suggested_Union_Proposed_Values	Suggested_Union_Proposed_cardinality	Suggested_union_mandatory/recommended/optional	If conditional, terms of condition stated here	Part of sunflower:	Link to terminology/ontology that defines item	Link to terminology/ontology that defines value(s)/valuelist(s)	Reasoning/explanation/evidence of/for suggestion		Concept: Exact match with	Concept: close match with	Example data (completely synthetic)		Link to terminology/ontology that defines the class	Class: Exact match with	Class: close match with				
Digital Resource	**Website**	Link to the website or repository (like GitHub) of the tool/software/database.	URL	0..n	recommended				NA													
Digital Resource	**Identifier**	bio.tools identifier for the digital resource.	bio.tools identifier	0..1	optional				bio.tools identifier													
Digital Resource	**Version**	The version of the tool/software/database used.	double	1..1	mandatory				NA													
Digital Resource	**Date used**	The date when the tool/software/database was last used.	xsd:dateTime	1..1	mandatory			https://www.w3.org/TR/prov-o/#endedAtTime	https://www.iso.org/iso-8601-date-and-time-format.html													
Digital Resource	**Settings**	Free text account of the settings used in the tool/software/database.	String	0..n	optional				NA				https://www.w3.org/TR/prov-o/#Plan									
Digital Resource	**Parameters**	Description of parameters used with the specified software. Copy the complete command line (all lines executed) used.	String	0..n	conditional	Mandatory if this Digital Reosurce is coming from Sequence: Alignment_software			NA													
Variant	**Variant_Type**	The category or type of variation or abnormality present in an amino acid or nucleic acid sequence.	SNVs, indels, SVs, CNVs, gene fusions, ... (to be extended)	0..n	recommended				SNVs, indels, SVs, CNVs, gene fusions, ... (to be extended)							http://www.sequenceontology.org/browser/current_release/term/SO:0001060							
Variant	**Variant_Origin**	A quality inhering in a variant by virtue of its origin.	somatic germline maternal paternal pedigree specific population specific de novo	0..1	recommended				somatic germline maternal paternal pedigree specific population specific de novo													
Variant	**Variant_representation**	The representation of the variant using HGVS nomenclature.	**String following HGVS nomenclature**	1..1	mandatory				https://hgvs-nomenclature.org/stable/													
Variant	**Clinical_Variant_Interpretation_criteria**	Internationally (e.g. ACMG, ESMO-ESCAT) criteria met for variant interpretation	**list of versions of ACMG, ESMO-ESCAT, others??**	0..n	optional				NA													
Variant	**Clinical_Variant_Interpretation_result**	Indicates result of clinical variant interpretation.	**benign, likely benign, VUS, likely pathogenic, pathogenic**	0..n	optional				**benign, likely benign, VUS, likely pathogenic, pathogenic**													
Variant	**Clinical_expert_panel_decision**	Decision by clinical expert panel concerning the variant interpretation	String	0..n	optional				NA													
Variant	**Applied_Criteria_of_Evidence**	A category which fits with categories provided by Expert panels or tools accepted in Clincial Practice. If such recommendations are not available the weighted categories provided by freely available tools would be acceptable.	**as listed in Tables 3 and 4 in the article with reference: https://www.acmg.net/docs/standards_guidelines_for_the_interpretation_of_sequence_variants.pdf**	0..n	optional				**as listed in Tables 3 and 4 in the article with reference: https://www.acmg.net/docs/standards_guidelines_for_the_interpretation_of_sequence_variants.pdf**													
Variant	**Clinical_Interpretation_Tool**	Identification of the tool used for clinical interpretation	**Digital Resource**	**0..n**	optional				NA													
Variant	**Variant_calling_software_deviation**	Identification of the software used for variant calling, if different from software stated in Sequence class.	**Digital Resource**	0..n	conditional	If panel of normals == T			NA													
Variant	**Variant_Annotation_tools_deviation**	Identification of the software used for variant annotation, if different from software stated in Sequence class.	**Digital Resource**	0..n	recommended	If variant calling is performed, software used for variant annotation stated here			NA	Conditional on file format: only applies to files with genomic variation such as VCF and gVCF.												
Variant	**Variant_Annotation_database_deviation**	Database and version used for variant annotation, if different from software stated in Sequence class.	**Digital Resource**	0..n	recommended	If variant calling is performed, database for variant annotation is to be stated		https://bioportal.bioontology.org/ontologies/FOAF/?p=classes&conceptid=http%3A%2F%2Fxmlns.com%2Ffoaf%2F0.1%2FOnlineAccount	NA	Conditional on file format: only applies to files with genomic variation such as VCF and gVCF.												
Variant	**Reported_to_patient**	Indication if the variant has been reported back to the patient, if different from software stated in Sequence class.	**boolean**	0..1	optional				NA	Added following this comment from Jenny Kovac: Should we add another item, declaring if variant has been reported back to the patient or not? There are cases where accidental findings, not related with primary referral reason, although clearly pathogenic, are not reported back to the patient. For example the age of a patient (adult onset disease variant identified in a child), the (un)actionability of the variant (some neurodegenerative disease associated variants). Mostly, these are cases where reporting would severely impact the quality of life of a patient and/or her/his family, without any direct clinical benefit.												
Target_Gene	**URI**	URI identifying the targeted gene.	URI to either HGNC, NCBI gene, OMIM, HPO or HGVS for variants	0..1	conditional	It is mandatory to provide either URI or label.			**URI to either HGNC, NCBI gene, OMIM or HGVS for variants**							https://bioportal.bioontology.org/ontologies/OBI?p=classes&conceptid=http%3A%2F%2Fpurl.jp%2Fbio%2F4%2Fid%2F2009080148990952B3		https://bioportal.bioontology.org/ontologies/OBI?p=classes&conceptid=http%3A%2F%2Fpurl.obolibrary.org%2Fobo%2FOBI_0001982	https://bioportal.bioontology.org/ontologies/MAXO?p=classes&conceptid=http%3A%2F%2Fpurl.obolibrary.org%2Fobo%2FMAXO_0001613			
Target_Gene	**Label**	Label of the target gene, if no URI can be provided.	String	0..1	conditional	It is mandatory to provide either URI or label.			NA													
Target_Gene	**Description**	Description of target gene.	String	0..1	optional				NA													