Clinical Annotations Help File

The set of files comprised of **clinical_annotations.tsv**, **clinical_ann_alleles.tsv**, **clinical_ann_evidence.tsv** and **clinical_ann_history.tsv** contain PharmGKB's clinical annotations and associated information. These annotations are manually created by the PharmGKB curators to provide an evidence-rated, genotype- or allele-based summary of the literature evidence annotated in PharmGKB for an association between a genetic variant and a drug. Please refer to the PharmGKB website for more information about clinical annotations, and how they are assigned a level of evidence based on scores.

Description of Files:

- **clinical_annotations.tsv**: Contains all of the meta-data about each clinical annotation.
- **clinical_ann_alleles.tsv**: Contains the genotype- or allele-based annotation text and CPIC-assigned allele function, if available.
- **clinical_ann_evidence.tsv**: Contains information about each supporting annotation (variant annotation, guideline annotation, label annotation) for every clinical annotation.
- **clinical_ann_history.tsv**: Contains the history of the clinical annotation, including the creation date and the dates of changes or updates to the annotation.
- **LICENSE.txt**: The PharmGKB license for using PharmGKB data, including clinical annotations.
- **CREATED_xxxx-xx-xx.txt**: This file indicates the date that all files in this group were created from the database.
- **README.pdf** file: This document.

A description of the fields in each file follows.

clinical_annotations.tsv:

- ***Clinical Annotation ID***: The unique PharmGKB ID number for the annotation.
- ***Variant/Alleles***: Variant rsID from dbSNP or the allele names.
- ***Gene***: HGNC gene symbol.
- ***Level of Evidence***: Levels 1A-4 with 1A being the highest level of evidence; more information found on PharmGKB.
- ***Level Override***: Description of whether the level of evidence assigned based on the clinical annotation score was changed by PharmGKB curators. Options: Yes (plus reason), No.
- ***Level Modifiers***: Description of extra information used when assigning level of evidence. Options: VIP Tier 1, rare variant.
- ***Score***: Clinical annotation score calculated from supporting annotations; more information found on PharmGKB.
- ***Phenotype Category***: Association phenotype. Options: toxicity, efficacy, dosage, metabolism/PK, PD, other.
- ***PMID Count***: The number of PMIDs with variant annotations used to support the clinical annotation.
- ***Evidence Count***: Number of annotations supporting the clinical annotation, including variant annotations, guideline annotations and drug label annotations.
- ***Drug(s)***: Drugs associated with the variant/allele.
- ***Phenotype(s)***: Phenotypes in the variant/allele-drug association. For example, if the association was found in patients with a particular phenotype (disease), or if the variant/allele-drug combination causes a particular phenotype.

- ***Latest History Date***: The date of creation of the clinical annotation or the last time it was updated.
- ***URL***: PharmGKB webpage for the clinical annotation.
- ***Specialty Population***: Description of a specialty population (e.g. 'Pediatric') in any supporting variant annotation.

Example row from clinical_ann_alleles.tsv file:

Clinical Annotation ID	Variant/Haplotypes	Gene	Level of Evidence	Level Override	Level Modifiers	Score	Phenotype Category	PMID Count	Evidence Count	Drug(s)	Pher
1447954390	rs75039782	CFTR	3	Yes: Level of evidence set to 3. Ataluren is a drug for the treatment of Duchenne muscular dystrophy caused by a nonsense mutation and not indicated in CF treatment.	Rare Variant; Tier 1 VIP	4	Other	2	2	ataluren	Cysti Fibro

---

clinical_ann_alleles.tsv:

- ***Clinical Annotation ID***: The unique PharmGKB ID for the annotation.
- ***Genotype/Allele***: The genotype or allele associated with the clinical phenotype in the next column.
- ***Annotation Text***: The clinical annotation for the given genotype or allele.
- ***Allele Function***: The CPIC allele function, if it has been assigned; more information found on PharmGKB.

Example rows from clinical_ann_alleles.tsv file:

Clinical Annotation ID	Genotype/Allele	Annotation Text	Allele Function
613979022	CC	May be less likely to have improved left ventricular ejection fraction after	

		carvedilol treatment.	
1183615480	*3	Patients carrying the CYP2D6*3 allele in combination with another no function allele may have decreased metabolism of carvedilol as compared to patients carrying two normal function alleles. This annotation only covers the pharmacokinetic relationship between CYP2D6 and carvedilol and does not include evidence about clinical outcomes. Other genetic and clinical factors may also influence carvedilol metabolism.	No function

---

clinical_ann_evidence.tsv:

- ***Clinical Annotation ID***: The unique PharmGKB ID for the annotation.
- ***Evidence ID***: The unique PharmGKB ID for the annotation supporting the clinical annotation, including variant annotations, guideline annotations and drug label annotations.
- ***Evidence Type***: The type of supporting annotation. Options: Variant Annotation (Drug), Variant Annotation (Phenotype), Variant Annotation (Functional Assay), Guideline Annotation, Label Annotation.
- ***Evidence URL***: PharmGKB webpage for the supporting annotation.
- ***Evidence PMID***: If the supporting annotation is a variant annotation, the PMID the variant annotation is based on; otherwise, blank.
- ***Evidence Summary***: Variant annotation text or description of the guideline or label.
- ***Study Parameter Used for Scoring***: The ID of the Study Parameters object used to determine score, only applicable to Variant Annotations
- ***Evidence Score***: The score of the supporting annotation.

Example row from clinical_ann_evidence.tsv file

Clinical Annotation ID	Evidence ID	Evidence Type	Evidence URL	Evidence PMID	Evidence Summary	Evi S
449717935	1449717924	Variant Drug Annotation	https://www.pharmgkb.org/variantAnnotation/1449717924	30136624	Genotype GG is associated with increased response to Opioid anesthetics, Other general anesthetics or volatile anesthetics as compared	3

	to genotypes AA + AG.

clinical_ann_history.tsv:

- ***Clinical Annotation ID***: The unique PharmGKB ID number for the annotation.
- ***Date (YYYY-MM-DD)***: The date of the history event.
- ***Type***: The type of the history event. Options: Create, Update, Note, Correction.
- ***Comment***: The comment entered by the PharmGKB curator describing the action taken on the clinical annotation; this field may be blank.

Example row from clinical_ann_history.tsv file:

Clinical Annotation ID	Date (YYYY-MM-DD)	Type	Comment
1450931822	2021-01-29	Update	Edited phenotype descriptions to include CPIC 'no recommendation'.

It is important to understand that clinical annotations are created from literature that has been curated by PharmGKB. There may be more literature in the public domain to support or contradict an association that is not in the PharmGKB database. PharmGKB does its best to manually curate high profile literature but does not contain curated literature from every domain-based journal, or all of PubMed. PharmGKB reviews evidence from curated literature in non-regular intervals and re-evaluates the evidence strength for each association as more literature becomes available.