Kevin, here are our proposals for the TBD or to be fixe terms LOINC terms.

1. The complicated one is for 81258-6. Which is now called allelic frequency and has a system of “patient”. In the V2 Spec which was published (2 years ago) and has been adopted by some vendors. The definition is as follows;

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 81258-6 | Allelic Frequency [NFR] | 2a | **0.47** | C | [0..1] | Reports the fraction of all of the reads at this genomic location that were represented by the given allele. For homozygotes it will be close to 1.0; for heterozygotes it will be close to 0.5. It can be a smaller number when there are mosaics or multiple chromosome, or mixtures of tumor cells and normal cells. |

A number of problems and tangles. First the description that is now in the LOINC database (as distinguished from what is in the V2 LRI definition, says it is the population frequency. However, the system was patient rather than population. Do not know how this was twisted. However, talked to the LOINC people and they proposed the following approach. (They may suggest some slight tweaks to the components to maintain internal consistency)

81258-6 Is specified as “trial” because it was created during lots of changes in the genomics proposal. They suggested locking this code to the definition now included in V2 Genetics standard because the only active use of this term is in that standard, and in that standard it is clearly defined as the sample variant, AND because the description saying it is a population based conflicts with its system which says it is a patient attribute. Because it is still tagged as a trial term, so it can be tweaked to be consistent.

We propose a new term for the population frequency. LOINC always included detailed information about changes in the release and in the RELMA program (With an icon that you click on to see)

Therefore, that is what this document proposes. We tweak the name of 81258-6 to include the word sample and Variant allelic frequency (and add a synonym of VAF) and use the definition-slightly edited that is tied to the Term in LRI V2 genomics reporting specification.

***Proposed Changes to 81258-6 now called Sample Variant Allelic Frequency use the description now recorded with this term in the LRI V2 specification for Allelic Frequency***

|  |  |
| --- | --- |
| Reference # |  |
| LOINC observation code | 81258-6 |
| LOINC Full name name | Sample Variant allelic frequency: NFr:Pt:^Patient:QN |
| Observation description | The fraction of all reads in a study sample at given genomic locus that identify the allele (variant) in question. For homozygotes it will be close to 1.0; for heterozygotes it will be close to 0.5. It can be less than 0.5 in the case of mosaics or multiple chromosome, or mixtures of tumor cells and normal cells. This measure is an attribute of the variant and applies when the method is a Next Generation Sequencing (NGS) or similar. Such methods provide many reads from the sample for each locus. |
| Reference Info/URL | Strom SP. Current practices and guidelines for clinical next-generation sequencing oncology testing. Cancer biology & medicine. 2016 Mar;13(1):3. |
| Component | Sample Variant allelic frequency |
| Component synonym | VAF |
| Property | NFr |
| Timing | Pt |
| System | ^Patient |
| Scale | Qn |
| Method |  |
| Answers |  |
| Units |  |
| Change Reason | Formerly called Allele Frequency (AF) |
| Related Names | Allelic fraction; novel allelic frequency |

1. We also created a new term for population frequency as shown below

***New term***

|  |  |
| --- | --- |
| Reference # |  |
| Local observation code | NA |
| LOINC Full name | Population allelic frequency:NFr:PT:^Population:QN |
| Observation description | The allelic frequency is the relative frequency (or fraction) of the given allele reported at the locus of interest in a given population. It is usually expressed as a number from 0 to 1. (https://www.sciencedirect.com/topics/neuroscience/allele-frequency) |
| Reference Info/URL |  |
| Component | Allelic frequency |
| Property | NFr |
| Timing | Pt |
| System | ^Population |
| Scale | Qn |
| Method |  |
| Answers |  |
| Units |  |
|  |  |

**New Term not previously discussed**…Requested by Bob F. to give some specificity for the source of the population allelic frequency. This will be a narrative text field to describe the source of the population frequency data.

|  |  |
| --- | --- |
| Reference # |  |
| Local observation code | NA |
| LOINC Full name | Description of the source of the population frequency data (will need help from LOINC for final name) |
| Observation description | Description of the source of the population frequency data |
| Reference Info/URL |  |
| Component | Description of the source of the population frequency data |
| Property | (Need help from LOINC) |
| Timing | Pt |
| System | ^Population |
| Scale | Narr |
| Method |  |
| Answers |  |
| Units |  |
|  |  |

1. Also created a term for coordinate system with three answers. It has long and rich description which I obtained from local NCBI experts. I passed this to Bob Dolan and he liked it. Think it is self-explanatory in practice the where the insertions go is tied to the kind of coordinate system and we don’t need to have another variable. Am grateful that bob pushed for simple. Though we end up with three rather than two answers.

***New TERM***

|  |  |
| --- | --- |
| Reference # |  |
|  |  |
| Full LOINC name | Genomic coordinate system:type:XXX:Nom |
| Observation description | A genomic coordinate is a position along a sequence. Coordinate systems can start counting from:  **0-base, interval counting**   * Used by GA4GH API, ClinGen data model, BED & BAM files, UCSC files, HLM 1.0 and NCBI’s SPDI conversion format * Akin to cursor positioning in modern text editors with cursor between characters * A bit unnatural for humans, easier for computers * Sequences have an inclusive start and exclusive end * Insertions always go in the interval between characters   **0-base character counting**   * Used by: Genbank database & ASN files * Sequences have an inclusive start, exclusive end * length = end – start * A bit unnatural for human reading, easier for computers * Sequences have an inclusive start and inclusive end * Insertions not defined because GenBank and ASN do not model insertions to records they contain. (historically ASN did internal to NCBI) SPDI was developed in part because of this inability to model insertions.   **1-base –character counting**   * Used by HGVS, VCF, NCBI’s ClinVar (uses HGVS), Genbank files, IPD-IMGT/HLA, UCSC genome browser (note different from UCSC file) * Akin to approach used in the earliest text processors * Length = start-end +1 * Sequences have inclusive start and inclusive end * More natural for humans * Insertions. Be careful in theory for 1-base, the insertion location could be defined in three ways. a) before position, b) after position, c) specify the di-nucleotide which where the insertion goes, but in practice there is only one way to do it per coordinate type.   **Distinctions for describing sub sequences with the three approaches with examples**  Consider ACG**TA**GTC as an example string and what nucleotides would be in the range 2-4   * 0-based interval counting (e.g..SPDI): GT * 0-based character counting (i.e. GenBank): GTA * 1-based character counting (i.e. VCF): CGT   **Distinctions regarding insertions with examples. Insertions can be tricky with character counting**  **Now consider that you want to describe an insertion between T&A (4th and 5th positions bolded and underlined) in the example):**   * 0 -based interval counting (i.e. SPDI) insert between 3 & 4 We know of no systems that support any other mechanism * 0 - based character counting (i.e. GenBank). Insertions are not supported in GenBank’s model. So, nothing to exemplify. * 1-based character counting. Insert between 4&5 (HGVS)- requires specification of the di-nucleotide between which the insertion goes. Insert After 4 (VCF)-does not require that specification but the result is the same. The above are the only cases that we know of. |
| Reference Info/URL |  |
| Component | Genomic coordinate system |
| Property | Type |
| Timing | Pt |
| System | XXX |
| Scale | Nom |
| Method | Molgen |
| Answers | • 0 -based interval counting  • 0 - based character counting  • 1-based character counting |
| Units |  |

1. New term not previously discussed… Kevin P. requested Level of Evidence

|  |  |
| --- | --- |
| Reference # |  |
| Local observation code | NA |
| LOINC Full name | ACMG Level of Evidence (Need help from LOINC to determine full name) |
| Observation description | Guidance from the American College of Medical Genetics and Genomics (ACMG) that recommends the use of standard terminology to describe variants identified in genes that cause Mendelian disorders and associated supporting evidence. |
| Reference Info/URL | [https://www.acmg.net/docs/standards\_guidelines\_for\_the\_interpretation\_of\_sequence\_variants.pdf](https://nam01.safelinks.protection.outlook.com/?url=https%3A%2F%2Fwww.acmg.net%2Fdocs%2Fstandards_guidelines_for_the_interpretation_of_sequence_variants.pdf&data=02%7C01%7CKevin.Power%40cerner.com%7C6c5729fa1f4d4cb2c29508d6c20fb70e%7Cfbc493a80d244454a815f4ca58e8c09d%7C0%7C1%7C636909770092320923&sdata=q6wKxhc14yfifV5rvC%2FOU4yb4XWlRmh8uc2nSZ1ZP%2Fg%3D&reserved=0) |
| Component | ACMG Level of Evidence |
| Property | Find |
| Timing | Pt |
| System | (Need help from LOINC) |
| Scale | Nom |
| Method | Reported |
| Answers | * Very strong evidence pathogenic * Strong evidence pathogenic * Moderate evidence pathogenic * Supporting evidence pathogenic * Supporting evidence benign * Strong evidence benign * Uncertain significance |
| Units |  |
|  |  |

1. New term not previously discussed… Kevin P. requested Level of Evidence

|  |  |
| --- | --- |
| Reference # |  |
| Local observation code | NA |
| LOINC Full name | CPIC Clinical Annotation Level of Evidence (Need help from LOINC to determine full name) |
| Observation description | Building from variant annotations in the knowledge base of PharmGKB, clinical annotations combine multiple-variant annotations into a single summary of the relevant variant-drug-phenotype association. Each clinical annotation is assigned a “level of evidence” score that is a measure of confidence in the association as determined by the PharmGKB curators. This score is based on several criteria, including replication of the association, P value (after correction for multiple-hypothesis testing, if applicable), and odds ratio, if available.  Levels of evidence diagram |
| Reference Info/URL | Pharmacogenomics knowledge for personalized medicine. Clinical pharmacology and therapeutics. 2012. Whirl-Carrillo M, McDonagh E M, Hebert J M, Gong L, Sangkuhl K, Thorn C F, Altman R B, Klein T E. [Article:22992668]  <https://www.pharmgkb.org/page/clinAnnLevels> |
| Component | CPIC Clinical Annotation Level of Evidence (Need help from LOINC to determine full name) |
| Property | Find |
| Timing | Pt |
| System | (Need help from LOINC) |
| Scale | Nom |
| Method | Reported |
| Answers | * Level 1A - Annotation for a variant-drug combination in a CPIC or medical society-endorsed PGx guideline, or implemented at a PGRN site or in another major health system. * Level 1B - Annotation for a variant-drug combination where the preponderance of evidence shows an association. The association must be replicated in more than one cohort with significant p-values, and preferably will have a strong effect size. * Level 2A - Annotation for a variant-drug combination that qualifies for level 2B where the variant is within a VIP (Very Important Pharmacogene) as defined by PharmGKB. The variants in level 2A are in known pharmacogenes, so functional significance is more likely. * Level 2B - Annotation for a variant-drug combination with moderate evidence of an association. The association must be replicated but there may be some studies that do not show statistical significance, and/or the effect size may be small. * Level 3 -Annotation for a variant-drug combination based on a single significant (not yet replicated) study or annotation for a variant-drug combination evaluated in multiple studies but lacking clear evidence of an association. * Level 4 - Annotation based on a case report, non-significant study or in vitro, molecular or functional assay evidence only. |
| Units |  |
|  |  |