Data Models (updated 10/21/2019)

- Publications (n=1112)
- Genes (n=127)
- Drugs (n=228)
- Guidelines (n=23)
- Gene-Drug Pairs (n=361)
- Functional Terms (n=17)
- Allele Translations (n=798)
- Allele Frequency (n=90,430)
- Gene Allele Functionality (n=781)
- Pre- and post-Test Alerts (n=339)
- Recommendations (n=23)
- Diplotype-Phenotype Mapping (n=24,634)

Allele Definition Files

- In DB
  - All except HLA-A and HLA-B
    - These require full sequence representation (or close to it) and never had “Allele Definition” files created with guidelines
    - Looking into options including specifying reference sequences or linking to HLA nomenclature
- Updated format; updated process to input to DB
- Updated versions on website

Allele Function Files

- In DB
  - CYP2B6
  - CYP2C9
  - CYP2C19
  - CYP2D6 updated for G2P project; copy number difficult
  - CYP3A5
  - NUDT15
  - TPMT
  - UGT1A1

- Updated format; updated process to input to DB
- Updated versions on the website as approved by author groups

Allele Function Files

- Problems
  - CYP4F2: has an allele (*2) with both increased and decreased function-no clinical function consensus
  - IFNL3 never had any allele function assignment – just recommendation based off of allele
  - SLCO1B1: has alleles with “possible” function that need to be cleared up first
  - DYPD: being finalized with author group
  - Non-standardized terms/no clinical function terms
    - CACNA1S
    - CFTR
    - G6PD
    - RYR1
    - HLA-A
    - HLA-B

Recommendation Tables (Table 1+2)

- CYP2B6/efavirenz in DB
- Issues:
  - More phenotypes in diplotype/phenotype file than recommendations
    - SLCO1B1/simvastatin
    - CYP3A5/tacrolimus
  - Phenotype groupings being updated – working with author groups now
    - CYP2C19 guidelines
    - CYP2D6 guidelines
  - Other
    - Warfarin
    - IFNL3/peginteferon written off of variants present, not function
    - G6PD hemi-zygous and female mosaicism
    - HLA's, CFTR, RYR1, CACNA1S – allele present so slightly different
- To do
  - CFTR, DPYD, thiopurines, anesthetics, CBZ, PHT, etc.

Diplotype-Phenotype

- Have tables created by hand by using manuscript table 1 + allele function file
- Should be able to re-create these automatically in DB if all content is there
- Provides a test for uploaded content AND a review of manual input
- Stuck at this point until function and recommendation issues are addressed

Gene-based CDS text/Flow chart

- Question if we need a flow chart for every gene
- CDS in:
  - CYP2B6
  - CYP3A5
  - NUDT15
  - TPMT
  - SLCO1B1
  - UGT1A1
- No CDS: IFNL3, CFTR, G6PD, CYP4F2, VKORC1
- Many footnotes being moved into comments column
- Some notes will be added to DB element for entire table

Drug-based CDS text/Flow chart

- Rose re-drew all flow charts to standardize
- CDS text in for all drugs for which a file was created
  - Had to split up many files created for guidelines with multiple drugs
- No CDS
  - Warfarin, Peginterferon alpha,

Frequency Files

- Varied greatly
- Good check of content and calculations
- In DB:
  - CYP2B6
  - CYP2C9
  - CYP2C19
  - CYP3A5
  - CYP4F2
  - HLA-A
  - HLA-B
  - NUDT15
  - SLCO1B1
  - TPMT
  - UGT1A1
  - VKORC1
  - CYP2D6
- Problems:
  - Changed population groupings – Rachel is updating these as needed
  - CYP2D6 had non-numbers entered (e.g. “>1000”)
  - IFNL3, G6PD, CFTR don’t have a file
- To do:
  - CACNA1S, RYR1, DPYD