

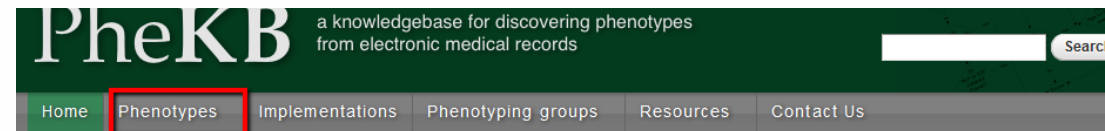


# PHENOTYPE CREATION THROUGH PHEKB.ORG

*Step-by-step Process Guide*

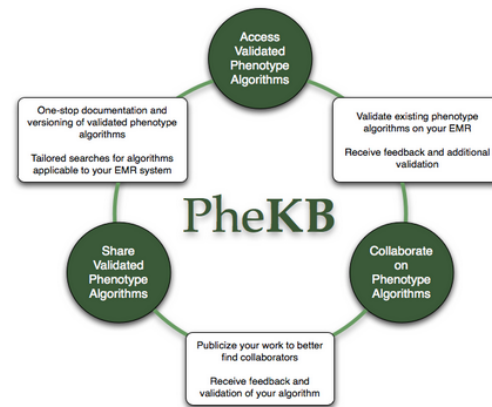
February 2, 2017

# Step 1: Hover over the “Phenotypes” tab on the main menu and select “Add Phenotype.”



## What is the Phenotype KnowledgeBase?

View Edit Revisions



Health Data is becoming an increasing important source for clinical and genomic research. Researchers create and iteratively refine algorithms using structured and unstructured data to better identify cohorts of subjects within the health data.

The Phenotype Knowledgebase website, PheKB, is a collaborative environment to building and validating electronic algorithms to identify characteristics of patients within health data. PheKB was functionally designed to enable such a workflow and has

purposefully integrated tools and standards that guide the user in efficiently navigating each of these stages from early stage development to public sharing and reuse. PheKB has tools to enable cross-site collaboration for algorithm development, validation, and sharing for reuse with confidence.

On PheKB you can: View existing algorithms; Enter or create new algorithms; Collaborate with others to create or review algorithms; View implementation details for existing algorithms

### Most Recent Phenotypes

	Sickle Cell Disease
	Consistent Care
	Peanut Allergy
	Liver cancer staging project
	Hypothyroidism

# Step 2: Add your phenotype by completing the Phenotype page creation form.

The screenshot shows the 'Create Phenotype' form with the following sections:

- Phenotype** (header)
- Flags**: No flags
- URL path settings**: Automatic alias
- Revision information**: New revision
- Comment settings**: Open
- Authoring information**: By Jacqueline Kirby
- Publishing options**: Published
- Title**: (empty text field)
- Body (Edit summary)**: (rich text editor)
- Text format**: Filtered HTML (dropdown menu)
- Status**: - Select a value - (dropdown menu)
- List on the Collaboration Phenotypes List**: (checkbox)

Annotations with green arrows point to the 'Status' dropdown and the 'List on the Collaboration Phenotypes List' checkbox.

Indicates Phenotype Status:  
In Development, Testing, Validation, Final

Indicates if Phenotype is posted on the  
Collaboration Tab

\*Note regarding **phenotype status**: Phenotype status marked as “Final” will be publicly accessible on the site. Status marked as “In Development,” “Testing,” or “Validation” will be privately accessible to only users who are logged into the site and part of either the phenotype’s “owner” or “view” group.

# Phenotype Creation Notes

- **Phenotype status:** Phenotype status selected as “Final” will be publicly accessible on the site. Status selected as “In Development,” “Testing,” or “Validation” will be privately accessible to only users who are logged into the site and part of either the phenotype’s “owner” or “view” group.
- **Owner vs. View Groups:** The *Owner* group or authoring group can edit the phenotype, comment on a phenotype, upload an implementation, and receive any data submission from sites implementing the phenotype. The *View* or collaborating group can subscribe to receive alerts when comments or modifications have been made to the phenotype, comment on a phenotype, upload an implementation, send a data submission to an authoring site based on site implementation.
- **Collaboration Phenotypes:** Checking the “List on the Collaboration Phenotypes List” option allows your phenotype to be publically listed on the [Collaboration Phenotypes page](#). Only the phenotype title, description, and contact information will be made available publicly to other researchers. This will enable researchers to contact Phenotype owners to collaborate on the phenotypes development.

# Step 3: Validate and post the data dictionary

- Once the phenotype has been created, upload the data dictionary through the Data Dictionary Tab of the phenotype.
- Results are categorized as “warnings” and “errors”. While it is recommended that you address all warnings and errors, in some cases it is understandable that warnings may be ignored.
- A Sample Data Dictionary and Corrected Data Dictionary can be found on the [PheKB Resources Tab](#).

The screenshot displays the PheKB website interface. At the top, the PheKB logo is visible with the tagline "a knowledgebase for discovering phenotypes from electronic medical records". A navigation menu includes Home, Phenotypes, Implementations, Phenotyping groups, Resources, and Contact Us. The main content area shows the phenotype "Major Adverse Cardiac Events while on Clopidogrel". Below the phenotype name, there are tabs for Phenotype, Edit, Data Dictionaries, Implementations/Datasets, and Revisions. The "Data Dictionaries" tab is selected, showing an "Upload a New Dictionary" section with a "Choose File" button and "Validate" and "Save" buttons. Below this, there is a section for an existing dictionary: "Dictionary - Clopidogrel\_Mace\_DD\_05072014.csv". This section includes a "Choose File" button, "Validate" and "Save Replacement" buttons, and a status message: "\* Dictionary: Clopidogrel\_Mace\_DD\_05072014.csv - Errors: 0 Warnings: 0".

# Data Validation

---

## Data Dictionary

No file chosen

### ▼ Data Dictionary Results -- sample-data-dictionary.csv

---

#### Warnings



##### File

---

- The SOURCE ID column is missing and should be included



##### Columns

---

- Column 4
  - The 4th column, SRCID, is not a standard column in data dictionaries



##### Rows

---

- Row 4
  - Row 4 appears to be blank and should be removed.