



PHEKB.ORG OVERVIEW

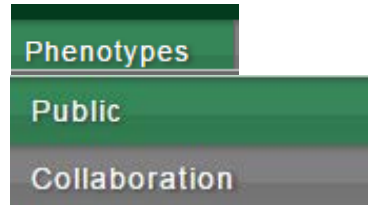
A Brief Introduction

February 2, 2017

PheKB: What is it and why would you use it?

- The Phenotype Knowledgebase (PheKB.org) is a workflow management system and learning center supporting the creation, validation, and dissemination of computable algorithms.
- PheKB was built as an outgrowth of lessons learned in eMERGE. We found this work requires a dynamic, social, and evolving tool for algorithm creation and sharing. It was created to make the iterative, collaborative work easier, allowing for:
 - *early assessment of feasibility for implementation*
 - *appropriate versioning*
 - *standardizing and reusing data elements for quality and efficiency*
 - *data quality and validation checks*
 - *timing and methods for disseminating the results.*

Public Repository



Public Phenotypes

Public Collaboration

Public phenotypes are believed to be complete and final by their authors. When you are logged in you can view and edit phenotypes in your groups that are non public and in various stages of development.

[Login To View Private Group Phenotypes](#)

Institution

- Beth Israel Deaconess Medical Center
- CHOP
- Cincinnati Children's Hospital Medical Center
- Columbia University
- Geisinger Health System
- Group Health and University of Washington
- Group Health Cooperative
- Marshfield Clinic Research Foundation
- Mayo Clinic

Leave blank for all. Otherwise, the first selected term will be the default instead of "Any".

Type of Phenotype

- Any -

Apply

Leave blank for all. Otherwise, the first selected term will be the default instead of "Any".

Title	Institution	Data Modalities and Methods Used	Owner Phenotyping Groups	View Groups	Has new content	Status	Type
Abdominal Aortic Aneurysm (AAA)	Geisinger Health System	CPT Codes, ICD 9 Codes, Vital Signs	eMERGE Geisinger Group	eMERGE Geisinger Group, eMERGE Phenotype WG		Final	Disease or Syndrome
ACE Inhibitor (ACE-I) induced cough	Vanderbilt University	CPT Codes, ICD 9 Codes, Medications, Natural Language Processing	eMERGE Vanderbilt Group	eMERGE Phenotype WG		Final	Drug Response - adverse effect or efficacy

Phenotypes that are finalized (status set to Final), and are ready to be shared publically. Often these phenotypes have associated publications to them.

Collaboration Phenotypes

Phenotypes

Public

Collaboration

Collaboration Phenotypes

Public

Collaboration



Collaboration phenotype authors in this list are open to engaging others in the development of the phenotype. The phenotype cannot be accessed until the author has shared access to the phenotype with you and you have logged into PheKB. Click the contact author email link to send an email if you are interested in working with the authors on a phenotype.

Type of Phenotype

- Any -

Apply

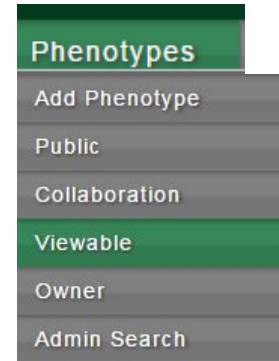
Leave blank for all. Otherwise, the first selected term will be the default instead of "Any".

Title	Institution	Data Modalities and Methods Used	Description	Status	Contact Author	Type of Phenotype
 Benign Prostatic Hyperplasia (BPH)	VUMC	CPT Codes, ICD 9 Codes, Laboratories, Medications, Natural Language Processing		Validated	Sarah Stallings	Disease or Syndrome
 Major Adverse Cardiac Events while on Clopidogrel	Vanderbilt University	CPT Codes, ICD 9 Codes, Medications	Patients with Major Adverse Cardiac Events (MACE) while on Clopidogrel. A MACE event can be either a revascularization event (bypass surgery or angioplasty/intracardiac stenting) or an acute myocardial infarction (AMI). A revascularization event is defined by the presence of a bypass, an angioplasty or a stent CPT	Testing	Wei-Qi Wei	Drug Response - adverse effect or efficacy

Collaboration Phenotypes – Public listing of phenotype on the Collaboration Phenotypes page. To list a phenotype on the collaboration page, check the box enabling this feature when editing a phenotype. Viewable on the Collaboration Tab:

- Phenotype Title
- Authoring Institution
- Data Modalities and Methods Used
- Description
- Status
- Contact Author with a link to contact author through PheKB
- Type Of Phenotype

A Collaborative Environment



Private Phenotypes-

Phenotypes with status of “In Development”, “Testing”, or “Validation” stages are not viewable publically and can only be viewed when a user is logged and the phenotype has been shared with the user via one of the two Collaborative groups:

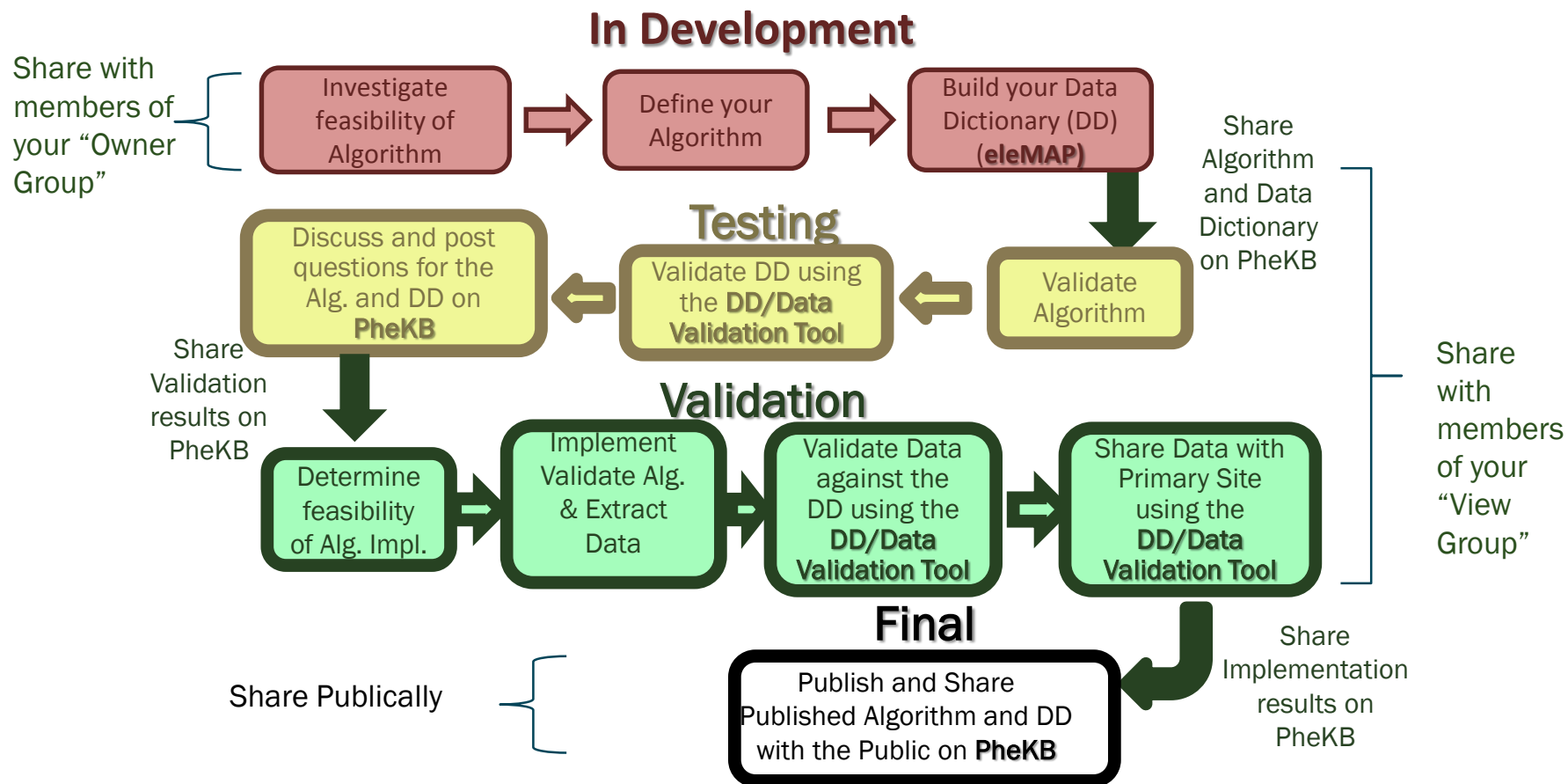
Owner Group Phenotypes- The Owner group or *authoring group*. The Owner group can

- edit the phenotype
- comment on a phenotype
- upload an implementation
- receive any data submission from sites implementing the phenotype

View Group Phenotypes – The View group or *collaborating group*. The View 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

Phenotype Workflow on PheKB



Other fields collected when creating a Phenotype

- Files- Phenotype and other associated files.
- Type Of Phenotype – Disease or Syndrome, Drug Response, Other Trait
- Data Modalities and Methods Used
- Network Associations
- Age
- Gender
- Race
- Ethnicity
- Authors and Contact Author
- Phenotyping Groups
- Institution
- URLs
- PubMed Reference

Validate Data Dictionary

PheKB a knowledgebase for discovering phenotypes from electronic medical records

Home Phenotypes Implementations Phenotyping groups Resources Contact Us

Phenotypes » Phenotype 300

Major Adverse Cardiac Events while on Clopidogrel

Phenotype Edit Data Dictionaries Implementations/Datasets Revisions

Upload a New Dictionary

Dictionary
Choose File No file chosen
Validate Save
You must validate before saving.

Dictionary - Clopidogrel_Mace_DD_05072014.csv

Clopidogrel_Mace_DD_05072014.csv (click to download)
Delete From Phenotype Datasets Using This

Replace file Clopidogrel_Mace_DD_05072014.csv
Choose File No file chosen
Validate Save Replacement

Dictionary: Clopidogrel_Mace_DD_05072014.csv - Errors: 0 Warnings: 0

Once the phenotype has been created, upload the data dictionary through the Data Dictionary Tab of the phenotype.

Note: The validation tool currently only accepts CSV files. If you have a data dictionary saved as an Excel document, you will need to convert it first.

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](#).

Posting results from Validation – Creating an Implementation

» Phenotypes

Implementations and Datasets For This Phenotype

Phenotype Data Dictionaries Implementations/Datasets

Upload a New Implementation

Implementation Details	Case PPV	Control PPV	Dataset/Dictionary	Dataset Validation
CDiff NU implementation - May 2012 chart review <i>Northwestern University</i> Cases: 0 Controls: 0 (Case, Control) Uploaded: 06/04/2012	0.902439	0.95	No datasets uploaded	
Group Health Implementation <i>Group Health Cooperative</i> Cases: 0 Controls: 0 (Case, Control) Uploaded: 06/06/2012	0.92	0.88	No datasets uploaded	
Vanderbilt implementation <i>Vanderbilt University</i> Cases: 0 Controls: 0 (Case, Control) Uploaded: 07/11/2012	1	1	No datasets uploaded	
Marshfield C-diff implementation <i>Marshfield Clinic Research Foundation</i> Cases: 21 Controls: 10 (Case, Control) Uploaded: 01/16/2014	1	1	No datasets uploaded	

When creating an implementation, you can include information regarding:

- Implementation Institution
- Details about site's implementation
- Citation where available
- Case and Control validation results for PPV and Sensitivity calculations
- Any associated implementation files

Note: Do not upload submission data when creating an implementation. That can be uploaded on the Dataset tab.

Data Set Validation and Submission

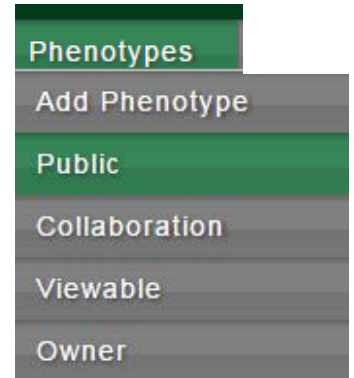
Once an implementation has been created. Users can upload data to an implementation to available, [validated data dictionaries](#).

A Sample Data file and Corrected Data file can be found on the [PheKB Resources Tab](#).



The screenshot displays the PheKB web application interface. At the top, the PheKB logo is accompanied by the tagline "a knowledgebase for discovering phenotypes from electronic medical records". The user is logged in as "Hello Jacqueline Kirby" with a "Logout" link. A search bar is visible in the top right corner. The navigation menu includes "Home", "Phenotypes", "Implementations", "Phenotyping groups", "Resources", and "Contact Us". The breadcrumb trail shows the path: Home » Phenotypes » Major Adverse Cardiac Events while on Clopidogrel » Implementations by Phenotype. The main content area is titled "MACE on Clopidogrel Vanderbilt Implementation" and features a "Datasets Validation" tab. Below the tab, there is a section for "Upload a New Dataset" which includes a "Select a dictionary" dropdown menu (currently showing "Clopidogrel_Mace_DD_05072014.csv"), a "Dataset" section with a "Choose File" button (indicating "No file chosen"), and a "Save and Validate" button. A note below the "Choose File" button states: "Upload a dataset file. Files larger than 97.75MB take time to validate and will be queued for validation."

Types of Sharing



Public Phenotypes – Phenotype set to Status = “Final”. These phenotype are ones ready to share broadly.

Collaboration Phenotypes – Public listing of phenotype on the Collaboration Phenotypes page.

Private Phenotypes-

Phenotypes with status of “In Development”, “Testing”, or “Validation” stages are not viewable publically and can only be viewed when a user is logged and the phenotype has been shared with the user via one of the two Collaborative groups:

- **Owner Group Phenotypes-** The Owner group or *authoring group*.
- **View Group Phenotypes** – The View group or *collaborating group*.