

Centralized Interactive Phenomics Resource (CIPHER): A Platform for Electronic Health Data-Based Phenomics Science

Systems Demonstration - Clinical Research Informatics

S56

Jacqueline Honerlaw, RN, MPH

VA Boston Healthcare System

#AMIA2023





I and my spouse/partner have no relevant relationships with commercial interests to disclose.

Learning Objectives

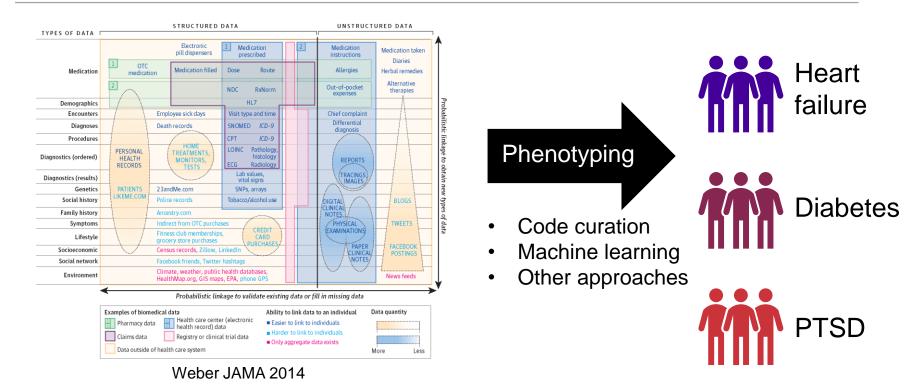


After participating in this session, the learner should be better able to leverage the Centralized Interactive Phenomics Resource (CIPHER) platform for electronic health records (EHR) based phenotyping by

- searching the library for phenotype definitions,
- contributing phenotypes, and
- browsing connected data visualization tools.

EHRs are a rich resource for clinical research and healthcare operations





Phenotype development has its challenges



Institutional knowledge



Computing resources

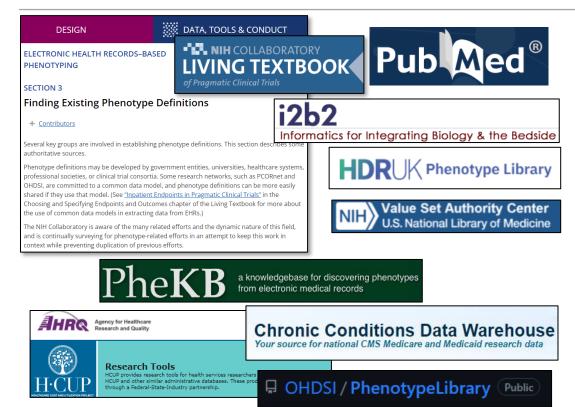




System specific data variability

Phenotype libraries allow reuse of existing definitions



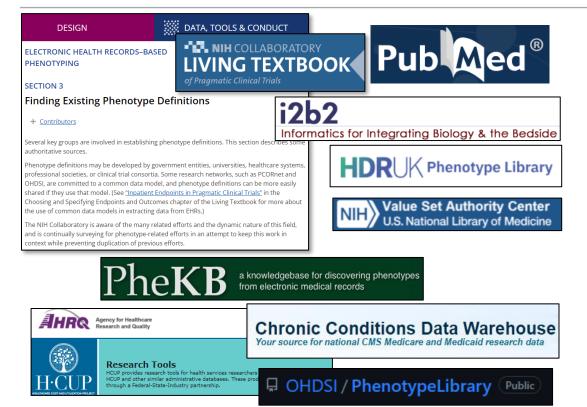


Benefits

- Reuse existing definitions
- Evaluate portability
- Save time and expedite reusability

There are opportunities for improvement and enhancement





Challenges

- Targeted user community
- Focused scope
- Metadata captured
- User features and capabilities

CIPHER: A public phenotype library developed in the VA by phenomics experts





Development Aims

- CIPHER metadata standard
- Scale content
 management
- Navigation and search
- Integrated visualization tools
- Health system agnostic

CIPHER's phenotype metadata standard



Phenotype Name

Phenotype Description Phenotype purpose and aims (E.g., to maximize sensitivity)

> Publication / Acknowledgment

Algorithm Description Logic applied to components to generate algorithm

Code Samples Programming code or code repository link

Population Population used to develop algorithm

Authors / Contacts

Data Sources Data source used to generate the phenotype

Data Classifications Disease, demographics, laboratory test, etc.

Role in Analysis Primary outcome/exposure, inclusion requirement, etc.

Related Disease Domain Cardiovascular, dermatology, genitourinary, etc.

Unique Identifier

Keywords

MESH Terms

Methods Used Rules based, machine learning, etc.

Context of Development Research, healthcare operations, etc.

Algorithm Components Medical Vocabularies (ICD, LOINC, etc.)

Algorithm Creation Date

Lab Units / Values

Algorithm Validation / Performance Metrics Sensitivity, specificity, etc.

Attachments Tables, figures, slides, etc.

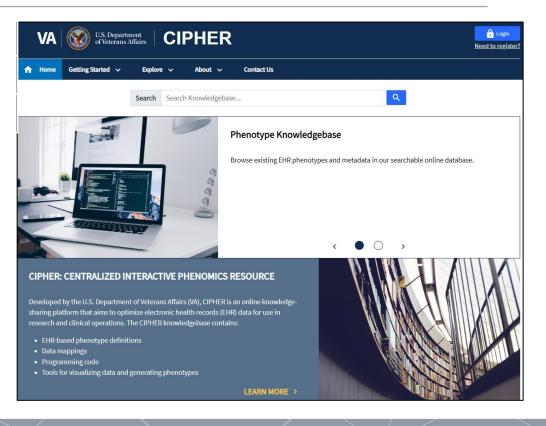
Legend

- Free Text
- Semi-controlled Authority
- Standard Vocabulary
- Empirical Value
-) Other

Honerlaw JAMIA 2023

CIPHER website components

- 1. Phenotype knowledgebase
- 2. Phenotype collection workflow
- 3. Data visualization tools





Searchable database of phenotype articles

1. Phenotype knowledgebase

- 2. Phenotype collection workflow
- 3. Data visualization tools

 ✓ CIPHER standard
 ✓ Unique phenotype article identifier
 ✓ Complex searching
 ✓ Change control

VA WS U.S. Department of Veterans Affairs	PHER ···
A Home Getting Started ∨ Explore ∨	About 🗸 Contact Us 🏟 Admin 🗸
X Clear Filters A Collapse	Search dementia
Data Classification	Search results for: dementia ×
Related Disease Domain	Sertoy Date algorithm created (descending) Titems per page: 10 T 1-2 xF3 LZ Z X X
🖹 Data Sources Used 🗸 🗸	Date algorithm created (descending) Items per page: 10 × 1-3 of 3 < < > >
Algorithm components 🗸	Alzheimer's Disease (MVP Cog Working Group) 😰
Role of phenotype in analysis	Author: MVP Cognitive Decline and Dementia During Aging Working Group, Million Veteran Program (MVP)
Date algorithm created	Algorithm Created: 12/22/2022
Author MVP Cognitive Decline and Dementia During Aging Working Group (3)	Logue MW, Miller MW, Sherva R, Zhang R, Harrington KM, Fonda JR, Merritt VC, Panizzon MS, Hauger RL, Wolf EJ, Neale Z, Gazlano JMF, Millon Veteran Program. Alzheimer's disease and related dementias among aging veterans: Examining gene-by-environment interactions with bootstraumatic stress disorder and traumatic brain injury. Alzheimers Dement. 2022 Dec 22. doi: 10.1002/alz.12870. Epub ahead of print. PMID: 36546606., MVP Cognitive Decline and Dementia During Aging Working Group, Million Veteran Program (MVP)
Million Veteran Program (MVP) (3)	Alzheimer's Disease, Non-specific Dementias (MVP Cog Working Group) 12
∑ Method used ✓	Author: MVP Cognitive Decline and Dementia During Aging Working Group, Million Veteran Program (MVP)
Publication	Algorithm Created: 01/01/2020
Yes 💌 🗙	Logue MW, Miller MW, Sherva R, Zhang R, Harrington KM, Fonda JR, Merritt VC, Panizzon MS, Hauger RL, Wolf EJ, Neale Z, Gaziano JM, Millon Veteran Program. Alzheimer's disease and related dementia s among aging veterans: Samining gene-by-environment interactions with post-traumatic stress disorder and traumatic brain injury. Alzheimers Dement.



Standardized collection of phenotype metadata



- 1. Phenotype knowledgebase
- 2. Phenotype collection workflow
- 3. Data visualization tools

 ✓ Validation against standard vocabularies
 ✓ CIPHER review
 ✓ Populate into knowledgebase

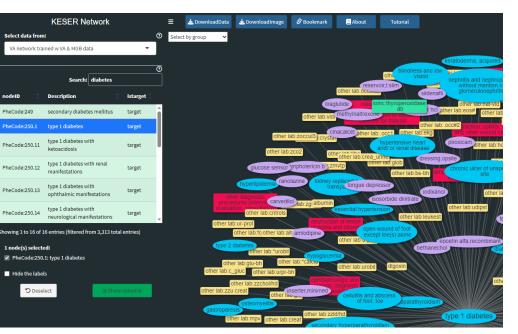
	н
Add Algorithm Component	
2 Enter Algorithm Component Codes	
ICD-9 Diagnostic Codes	
How would like to add the information?	
O Search for a code by its name or description	
Perform a wildcard search by using % or *	
$O \sum_{\rm special characters are allowed.}^{\rm Enter codes directly. You may provide multiple codes separated by a comma. No other$	
434.%	
✓ 434.0 - CEREBRAL THROMBOSIS	
434.00 - CEREBRAL THROMBOSIS W/O MENTION OF CEREBRAL INFARCTION	
✓ 434.01 - CEREBRAL THROMBOSIS W/ CEREBRAL INFARCTION	
434.10 - CEREBRAL EMBOLISM W/O MENTION OF CEREBRAL INFARCTION	
✓ 434.1 - CEREBRAL EMBOLISM	
433.01 🛢 433.11 🛢 433.21 🛢 433.31 🛢	
433.81 433.91 434.01 434.11	
434.91 1 434.0 1 434.1 1 434.9 1	
	. 1
OK Cancel	

AMIA 2023 Annual Symposium | amia.org

Tools connected to phenotype definition knowledgebase

- 1. Phenotype knowledgebase
- 2. Phenotype collection workflow
- 3. Data visualization tools

- ✓ ICD Hierarchy: View code mappings
 ✓ KESER: Aid phenotype development
- ✓ Link to knowledgebase

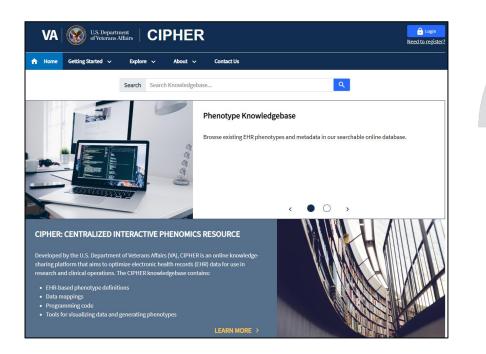


Hong NPJ Digit Med 2021



CIPHER demonstration





CIPHER

CENTRALIZED INTERACTIVE PHENOMICS RESOURCE



https://phenomics.va.ornl.gov/

CIPHER's user base continues to grow



Current Users





U.S. Department of Veterans Affairs



🛄 Mass General Brigham





Total Phenotypes 4,200+ Published 3,900+

Future Collaborations





TERS FOR DISEASE CONTROL AND PREVENTION



Agency for Healthcare **Research and Quality**







U.S. Department of Defense





Strengths and limitations of CIPHER



Strengths

- Standard metadata collection
- Healthcare system agnostic
- Validation of standard vocabularies
- Integrated data visualization tools
- Structure is designed for expansion

Limitations

- Licensed data elements may have restrictions (UMLS)
- No linkage to patient level data

Future directions and next steps



Future directions

- Expand database content
- Integrate additional tools
- Expand partners and contributors
- Collect user feedback

How you can use CIPHER

- Browse phenotypes and tools
- Contribute phenotypes and tools
- Share user feedback and ideas

Acknowledgements

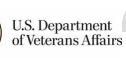
- Kelly Cho
- Anne Ho
- Francesca Fontin
- Ashley Galloway
- Jeff Gosian
- Monika Maripuri
- Michael Murray
- Rahul Sangar
- Joanne Sordillo
- Vidisha Tanukonda
- Edward Zielinski

- Suma Muralidhar
- Mike Gaziano
- Rachel Ramoni
- Katherine Liao
- Tianxi Cai
- David Gagnon
- Stacey Whitbourne
- SuChun Cheng
- Vidul Panickan
- Andy Zimolzak
- John Russo

- Laura Davies
- David Heise
- Keith Connatser
- Chad Steed
- Anuj Kapadia

















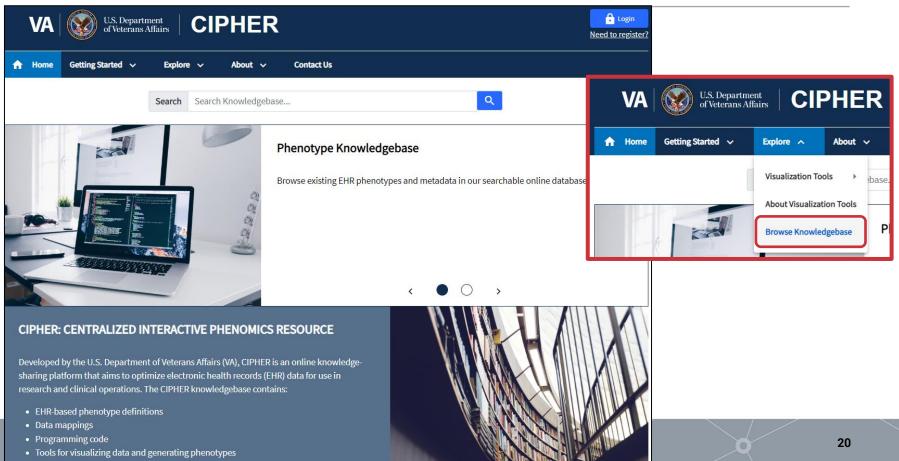
https://phenomics.va.ornl.gov/

Thank you!

Email me at: Jacqueline.Honerlaw@va.gov

Search for a phenotype definition





Search for a phenotype definition



Login

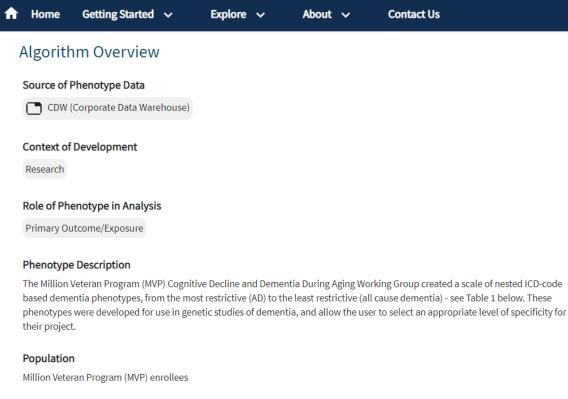
VA U.S. Departm of Veterans A	ent ffairs CIF	PHER Need to rep
↑ Home Getting Started ∨	Explore 🗸	About V Contact Us
🗙 Clear Filters 🥆 Collapse	<u>.</u>	Search dementia
Data Classification	~	Search results for: dementia ×
Related Disease Domain	~	Sort by Search relevance Items per page: 10 • 1-3 of 3
Data Sources Used	~	
Algorithm components	~	Alzheimer's Disease, Non-specific Dementias (MVP Cog Working Group) [2]
Role of phenotype in analysis	~	Author: MVP Cognitive Decline and Dementia During Aging Working Group, Million Veteran Program (MVP)
Date algorithm created	~	Algorithm Created: 01/01/2020
Author	~	Logue MW, Miller MW, Sherva R, Zhang R, Harrington KM, Fonda JR, Merritt VC, Panizzon MS, Hauger RL, Wolf EJ, Neale Z, Gaziano JM; Million Veteran Program. Alzheimer's disease and related dementias among aging veterans: Examining
∑ Method used	~	gene-by-environment interactions with post-traumatic stress disorder and traumatic brain injury. Alzheimers Dement. 2022 Dec 22. doi: 10.1002/alz.12870. Epub ahead of print. PMID: 36546606., MVP Cognitive Decline and Dementia During Aging Working Group, Alzheimer's Disease, Non-specific Dementias (MVP Cog Working Group), Alzheimerâ® [®] s Disease,
Publication	~	Non-specific Dementias_attachment.docx
Algorithm code	~	Alzheimer's Disease, Related Dementias (MVP Cog Working Group) Z
⊐ ¥ Validated	^	Author: MVP Cognitive Decline and Dementia During Aging Working Group, Million Veteran Program (MVP)
Yes	• ×	Algorithm Created: 01/01/2020

Review definition: Basic information



★ Home Getting Started ∨ Explore ∨ About ∨	Contact Us
Basic Information and Contact	^
General Phenotype	Data Classification(s)
Alzheimer's Disease, Non-specific Dementias (MVP Cog Working Group)	Diseases
Abbreviations and Keywords	Related Disease
AD+	Mental/ Behavioral Health Neurology
Publication	Author(s)
Logue MW, Miller MW, Sherva R, Zhang R, Harrington KM, Fonda JR, Merritt VC, Pan Million Veteran Program. Alzheimer's disease and related dementias among aging	
interactions with post-traumatic stress disorder and traumatic brain injury. Alzheir Epub ahead of print. PMID: 36546606.	
· · ·	MVP Cognitive Decline and Dementia During Aging Working Group <u>mark.logue@va.gov</u>
	Acknowledgement
	Mark W. Logue, Richard Hauger, Victoria Merritt, and Matthew Panizzon. Phenotypes developed by the MVP Cognitive Decline and Dementia During Aging Working Group, with support from VA Grants BX004192 (MWL), BX005749 (MWL), I01 CX001727 (RH), and IK2 CX001952 (VM).

Review definition: Algorithm overview





Date Algorithm Created

01/01/2020

Review definition: Algorithm components



Home	Getting Started	~	Explore	~	About	~	Contact Us
Algorith	nm Compone	ents					
Method Us	sed						
Rules-Base	ed (i.e., only structure	ed data wer	re used)				
To quality a	Description is a case requires the	presence of	two or mor	re ICD code:	S		
Algorithm	Components						
ICD-9 Diag	gnostic Codes (8)						
290.0, 290).20, 290.21, 290.3, 29	4.20, 294.21	1, 294.8, 331	0			
ICD-10 Dia	agnostic Codes (5)						
F03.90, F0)3.91, G30.1, G30.8, G	30.9					
(1) Program	mming Code Other						
/*****	*****	MVP Phen	otype Cod	e for De	mentia.	ADRD. AD	Plus and Stricter AD

Project Name: Alzheimer's Phenotyping in MVP

Review definition: Validation & additional information

Home Getting Started 🗸 Explore 🗸 About 🗸 Cor	ntact Us			
Validation		^		
Validated Yes	Additional Attachments	Information		^
Description of Validation	Туре		Description	
We have performed a chart review of n=39 "difficult cases". That is, MVP participants with eit dementia, or b) subjects with low to moderate AD case probability according to the Multimo (Liao et al. Am Med Inform Assoc, 2019). These were reviewed in tandem by Drs. Hauger and	Alzheimer's Dise	ase, Non-specific Dementias	Additional Information - ICD codes, Performance	e of MAP
classified as "Likely not", "Possible", or "Likely" cases of MCI, AD, and dementia. Chart review classifications according to the MAP algorithm and according to our ICD-code based definitio 2). The performance of the ICD code set in shown in "Algorithm Performance Measures". The our ICD code based phenotypes in genetic studies by testing their association with the APOE factor) in European-descent MVP subjects (Table 3). Performance Metric	ons of MCI, AD, Al authors also eva	DRD, and dementia (Table luated the suitability of		
Sensitivity		0.82		
Specificity		0.71		
Negative Predictive Value (NPV)		0.75		
Positive Predictive Value (PPV)		0.78		
Area Under the ROC Curve (AUC)		N/A		
Карра				25



🔒 Home 🛛 Getting Started 🗸 🛛 Explore 🗸 About 🗸 Contact Us 🏟 Admin 🗸

Contributing Phenotypes

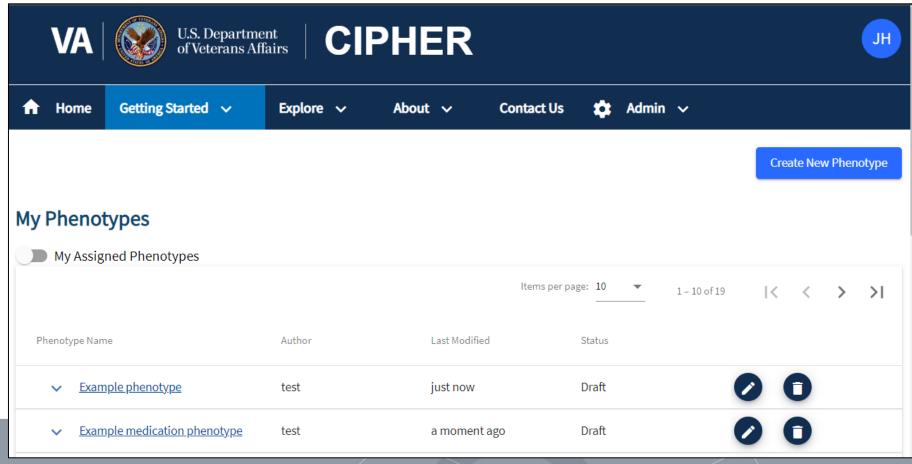
We have an easy-to-use process and resources for contributing your phenotype to CIPHER.

You may start by visiting our <u>How to Use CIPHER</u> page, which gives an overview of phenotyping, how phenotypes are used, benefits of contributing your phenotype to CIPHER, and general tips for using the CIPHER website.

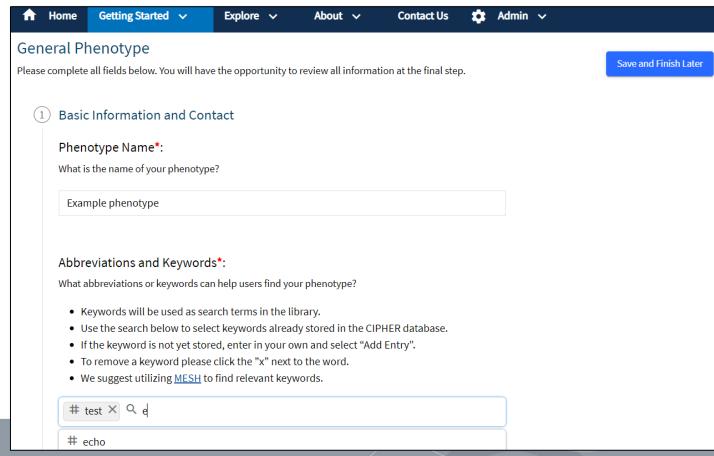
Next steps to contribute a phenotype:

- 1. Create a user account. Navigate to the top right-hand corner of the screen and choose to register a new account. Please enter a valid email that you have access to. You will need this to verify your account.
- 2. Complete the Phenotype Entry Form. Open the phenotype entry form by clicking <u>here</u>. You can also navigate to the form directly via the navigation bar at the top of the screen, under the Getting Started tab. You may save your work in progress and return to finish entering your phenotype at any time. To resume editing, please click on "My Phenotypes" under the "Getting Started" tab on the navigation bar.
 - a. Choose the type of phenotype you would like to submit (general, lab, or medication).
 - b. Follow the instructions in the online wizard to complete the form.
- 3. Submit your phenotype. Once you have entered in all your phenotype details, you can review your entry and submit for review.
 - a. Once your phenotype is submitted it will be reviewed by the CIPHER team. If our team has any questions regarding the submission, we will reach out to the contact listed on the entry form. Once the phenotype entry is finalized by CIPHER it will be searchable within our <u>Phenotype Knowledgebase</u>.
 - b. You can check the status of your submission by navigating to "My Phenotypes" on the navigation bar under the Getting Started tab.
 - c. Should you have any questions about your submission in the meantime, please contact <u>CIPHER@va.gov</u>.









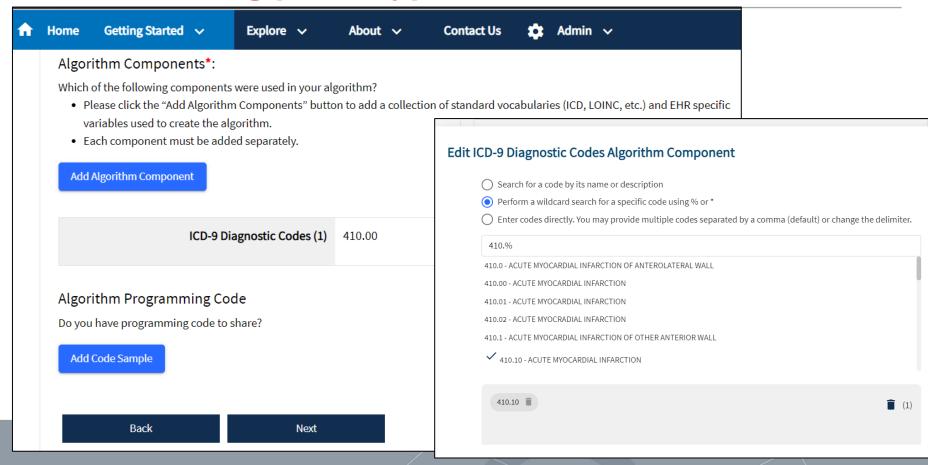


♠	Home	Getting Started 🗸	Explore 🗸	About 🗸
	🖉 Basic	Information and Con	tact	
	2) Algori	ithm Overview		Context of For what purp
	Sourc	e of Phenotype Data*:		Research
	What da	ata sources were used to gen ta	erate the phenotype	e 🖌 Healthca
		CDW (Corporate Data War	ehouse)	Clinical C
] CSDR (COVID-19 Shared D] DoD (Department of Defe		🗌 Other
] MVP (Million Veteran Prog] OMOP (Observational Mee	-	Role of Phe
	_	Other (VA Data Source)	uical Outcomes Pal	Primary (
		/A Data CMS (Medicare & Medicaio	d)	Secondar
] NDI (National Death Index		Comorbi
		Other		Other

🔒 Home	Getting Started 🗸	Explore	~	About	~	Contact Us	\$
	rithm Components						
	tunn components						
Meth	od Used*:						
What l	ogic was applied to the alg	orithm compo	nents to c	create this	pheno	type? (Select all that	t apply)
🗹 Ru	lles-Based (i.e., only struc	tured data w	ere used)				
🗌 Ma	achine learning: Supervis	ed					
🗌 Ma	achine learning: Semi-Sup	ervised					
🗌 Ma	achine learning: Unsuper	vised					
🗌 Ma	achine learning: Other ma	chine <mark>l</mark> earnin	ig approa	ich			
🗌 Ot	her						



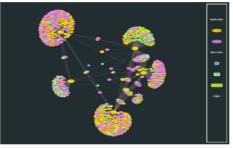




Visualization tools: KESER



VA U.S. Department of Veterans Affairs Affairs Getting Started V	Explore 🔨 About 🗸	Contact Us 🌼 Admin 🗸
Billing codes,	r Visualization Tools	es (ICD) codes, are often leveraged to define a patient's not always organized meaningfully for the purpose of high-
throughput p 10 billing cod	About Visualization Tools	the use of ICD codes for research by regrouping ICD-9 and ICD-
	Browse Knowledgebase	
	Million Veteran Program (MVP)	



Knowledge Extraction via Sparse Embedding Regression Network

Infer relatedness among diseases, treatment, procedures and laboratory measurements by creating a visual, interactive knowledge map.

Overview

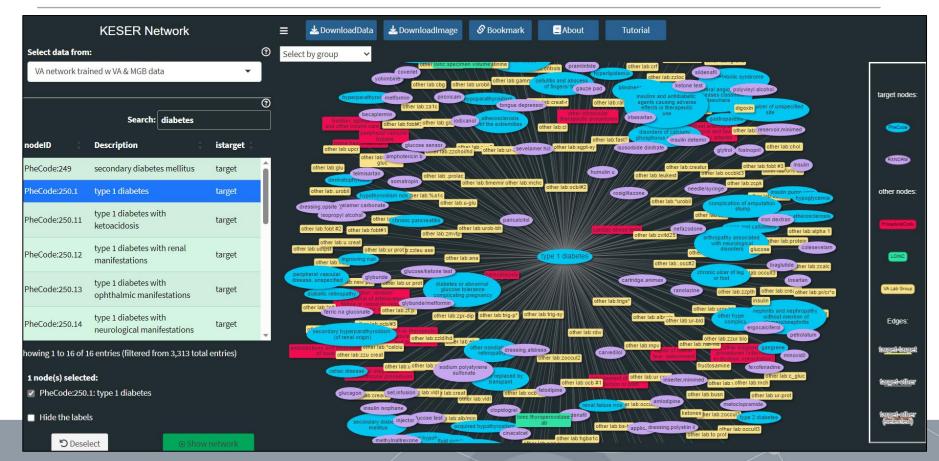
Uses

Author & Citation

The increasing availability of EHR systems has created enormous potential for translational research. However, it is difficult to know all the relevant codes related to a phenotype due to the large number of codes available. Codified concepts include all ICD diagnosis codes (rolled up to PheCodes), medications (mapped to RxNorm at ingredient level), procedures (rolled up to clinical classification system (CCS) procedure codes), as well as laboratory test results (mapped Logical Observation Identifier Names and Codes (LOINC), short names at the U.S. Department of Veterans Affairs (VA), or local lab codes).

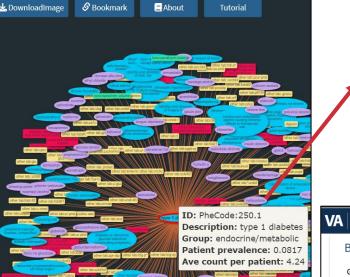
Visualization tools: KESER

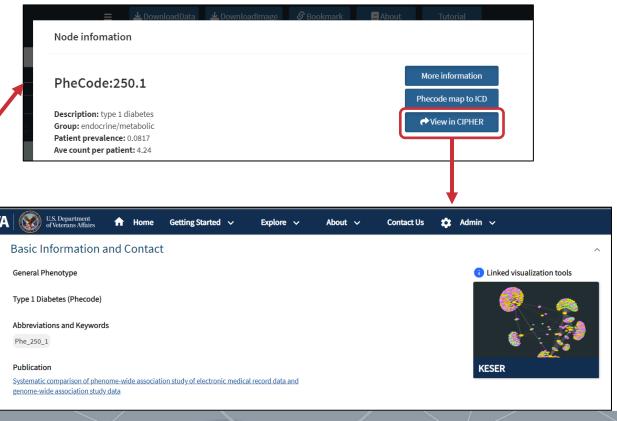




Visualization tools: KESER







Visualization tools: ICD Hierarchy



U.S. Department of Veterans Affairs Home Getting Started 🗸 Explore ~ About 🗸 Contact Us Admin 🗸 ICD Hierarchy Visualize mappings from Phecode to ICD-9 and ICD-10 codes. Overview Author & Citation Uses Electronic health record (EHR)-based studies offer several advantages in research: they are cost efficient, allow for large scale longitudinal analyses, and provide the potential to analyze hundreds of human diseases, drug responses, and many observable clinical traits. Billing codes, or International Classification of Diseases (ICD) codes, are often leveraged to define a patient's phenotype in EHR-based studies. However, these are not always organized meaningfully for the purpose of highthroughput phenotypic analyses. PheCodes facilitate the use of ICD codes for research by regrouping ICD-9 and ICD-10 billing codes into clinically relevant phenotypes.

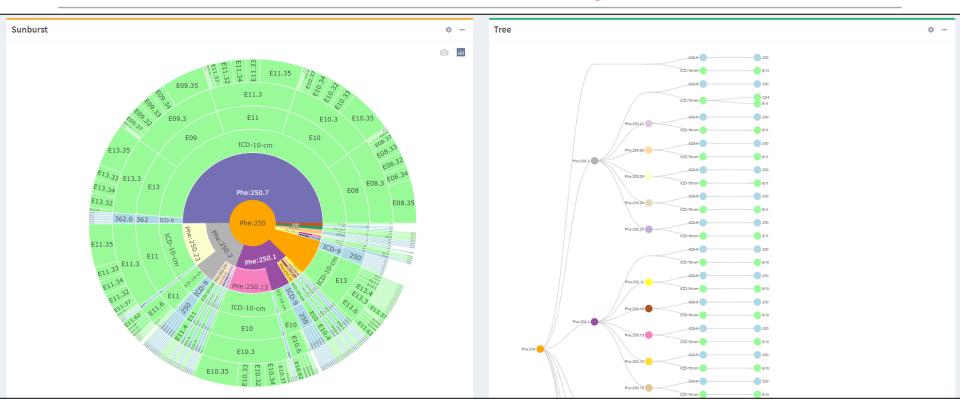
Visualization tools: ICD Hierarchy



Phecode	Phenotype	ICD version	ICD code	ICD Description	
250 🛞	All	All	All	All	ICD-10-cm
250.1	Type 1 diabetes	ICD-9	250.01	type I diabetes mellitus [insulin dependent type] [IDDM] [juvenile type], not stated as uncontrolled, without mention of complication	 ICD-9 Phe:250.23 Phe:250 Phe:250.24
250.1	Type 1 diabetes	ICD-9	250.03	type I diabetes mellitus [juvenile type], uncontrolled, without mention of complication	Phe:250.1
250.11	Type 1 diabetes with ketoacidosis	ICD-9	250.11	type I diabetes mellitus [insulin dependent type] [IDDM] [juvenile type], not stated as uncontrolled, with ketoacidosis	 Phe:250.11 Phe:250.3 Phe:250.12 Phe:250.4
250.11	Type 1 diabetes with ketoacidosis	ICD-9	250.13	type I diabetes mellitus [juvenile type], uncontrolled, with ketoacidosis	 Phe:250.13 Phe:250.14 Phe:250.14 Phe:250.42
250.1	Type 1 diabetes	ICD-9	250.21	Diabetes mellitus, type I [insulin dependent type] [IDDM] [juvenile type] with hyperosmolarity, not stated as uncontrolled	 Phe:250.15 Phe:250.2 Phe:250.6
250.1	Type 1 diabetes	ICD-9	250.23	Diabetes mellitus, type I [juvenile type] with hyperosmolarity, uncontrolled	Phe:250.21 Phe:250.7
250.1	Type 1 diabetes	ICD-9	250.31	Diabetes mellitus, type I [insulin dependent type] [IDDM] [juvenile type], not stated as uncontrolled	
250.1	Type 1 diabetes	ICD-9	250.33	Diabetes mellitus with other coma, type I [juvenile type], uncontrolled	
				Previous 1 2 3 4 5 17 Next	-
unburst				tree	

Visualization tools: ICD Hierarchy





Guidelines for using CIPHER



