JOHNS HOPKINS
MEDICINE

Innovations in use of biomedical informatics, electronic medical records and other Big Data for improving health systems

Casey Lynnette Overby, PhD
Assistant Professor, Division of General Internal Medicine
Division of Health Sciences Informatics
Malone Center for Engineering in Health Care
Johns Hopkins University

Behavioral, Social & Systems Science Translational Research Community
May 5th, 2017

About me

- My view
  - Biomedical informatics is the analysis, management, and use of knowledge, information and data ("Big Data") in the domain of biomedicine and health. *(Kulikowski et al. JAMIA 2012)*
  - Public health genetics provides context for genomic discoveries including complex ethical, legal, policy and social issues

- Involved in two projects
  - NHGRI-funded electronic medical records and genomics (eMERGE) Network (2011 – current)
  - NCATS Biomedical Data Translator Program (2016 – current)

- Strategic planning panel on NLM's role in supporting the publics health (April 2017)

Challenges to leveraging current innovations in using big data to improve health systems

JOHNS HOPKINS MEDICINE

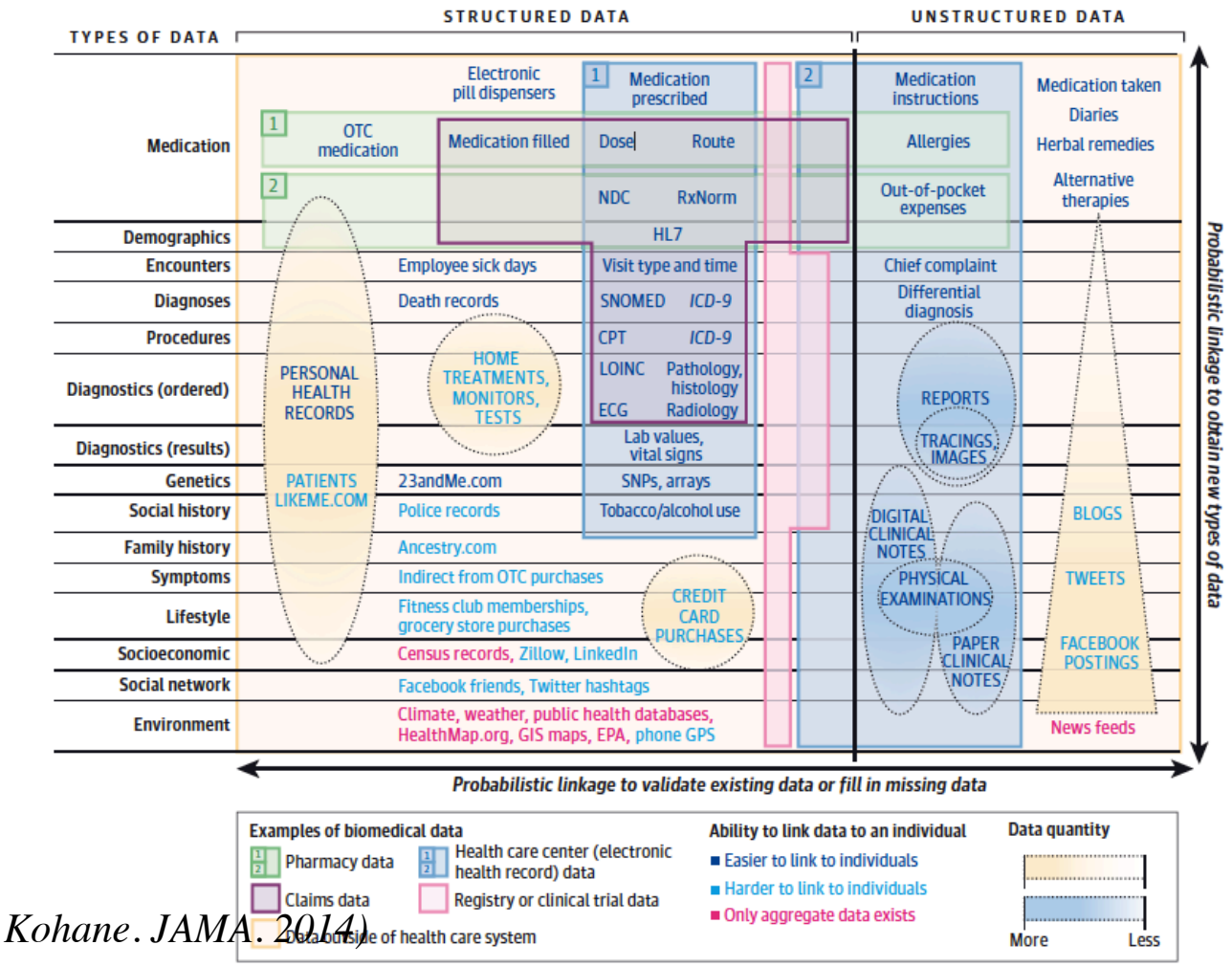
- Challenge #1: How can we decipher the meaning of data collected from various sources?

- Challenge #2: How can we deliver new evidence from (big) data analyses in an effective way?

New "omics" technologies, sensors, and social networks platforms provide access to new forms of population health data that can be combined with data from healthcare settings to improve how we deliver healthcare

JOHNS HOPKINS MEDICINE

Figure. The Tapestry of Potentially High-Value Information Sources That May be Linked to an Individual for Use in Health Care



(Weber, Mandl, Kohane. JAMA. 2014)

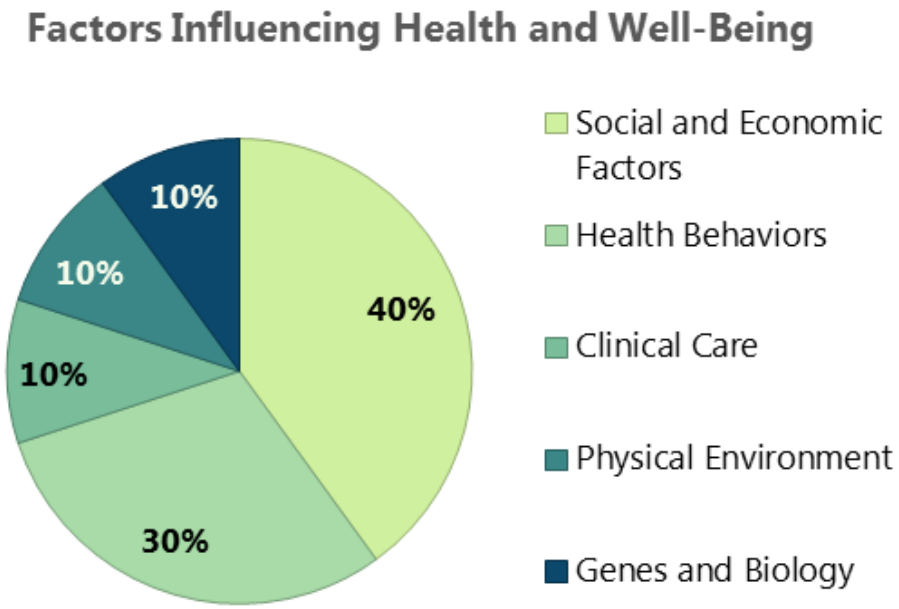
New measurement sources

Potential control variable:
Age, race, sex, genetic factors…

Wellness & exposure measures

Outcome measures

Health outcomes
Length of life
Quality of life
…



(Source: http://www.health.state.mn.us/divs/opi/gov/chsadmin/images/_factors_rev.png)

Predictive algorithms… beware of G.I.G.O. (the data is not the problem)

Input

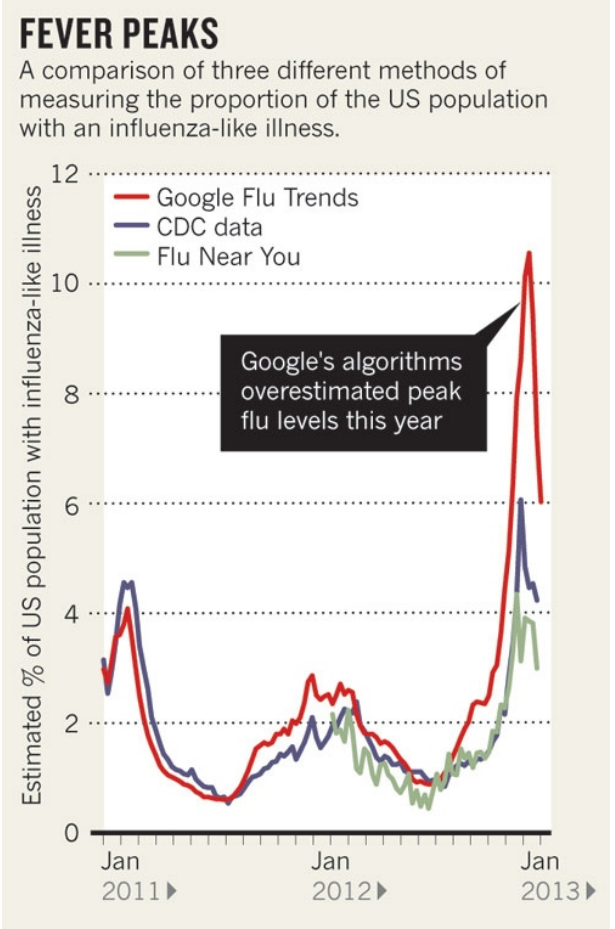
"Big Data analytics"

Output (Meaningful?)

e.g., 50% chance of rain

e.g., PPV <20%

e.g., Google Flu Trends
(Butler et al. Nature 2013)



How can we decipher the meaning of data collected from various sources?

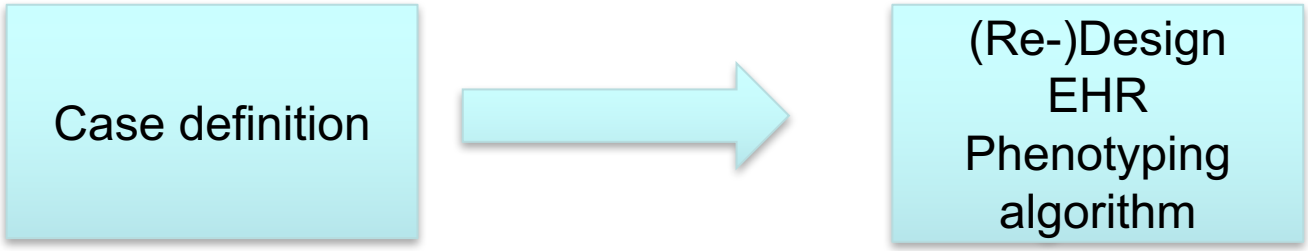
JOHNS HOPKINS MEDICINE

- Capturing the value of data from multiple sources for a specific context
- Biomedical informatics strives to link knowledge across the entirety of biomedicine
- EHR phenotyping is one approach that requires using data from multiple sources

Overview of EHR phenotyping process

Case definition

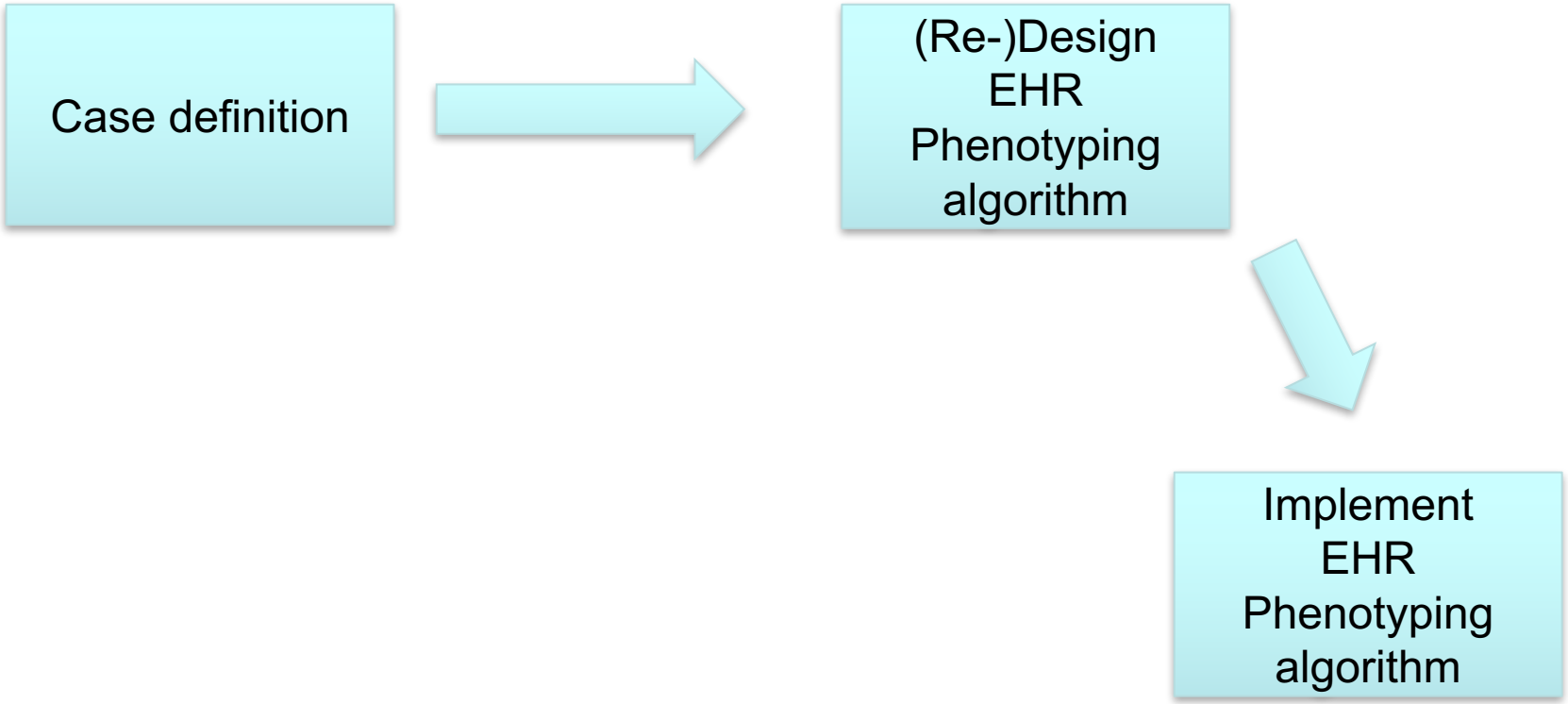
Overview of EHR phenotyping process



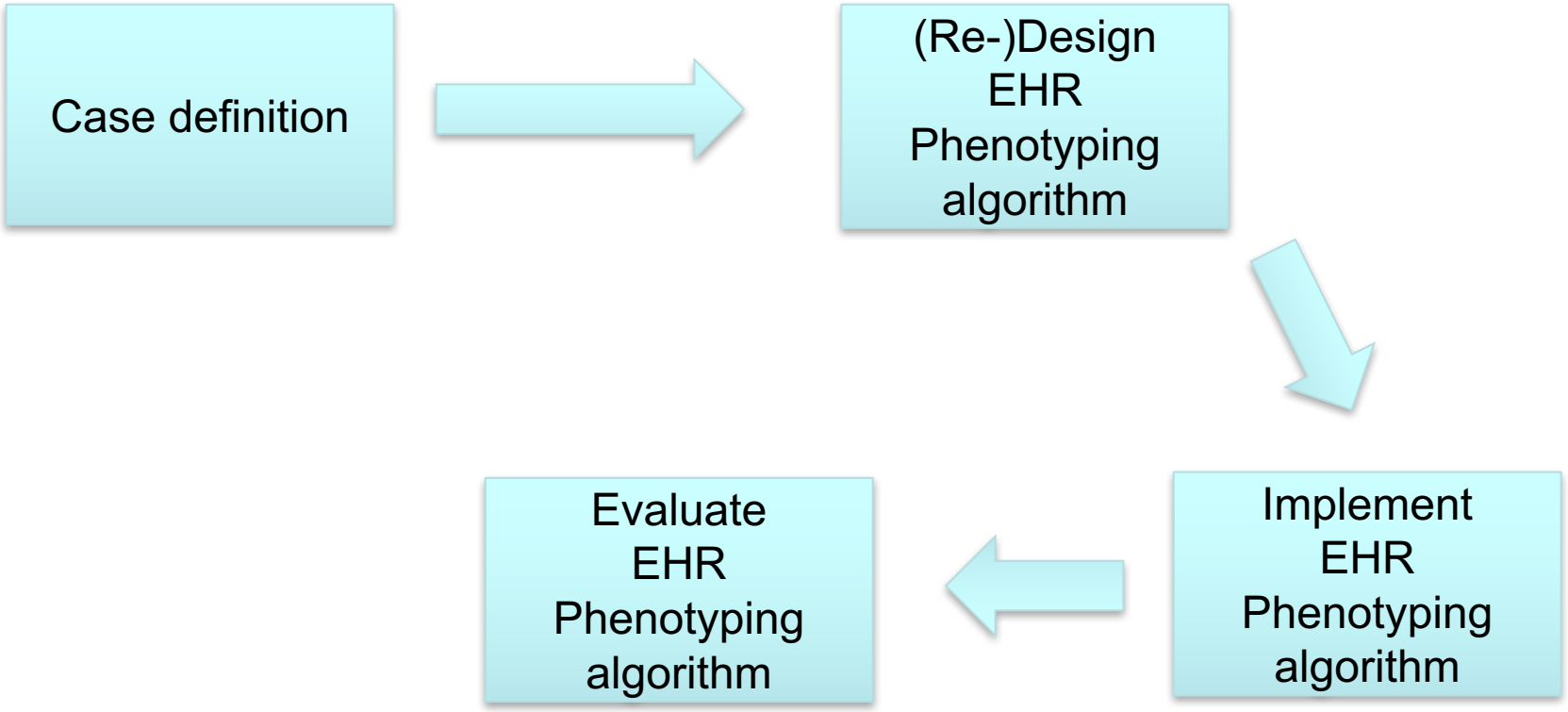
e.g., liver injury

e.g., ICD-9 codes for acute liver injury, Decreased liver function lab

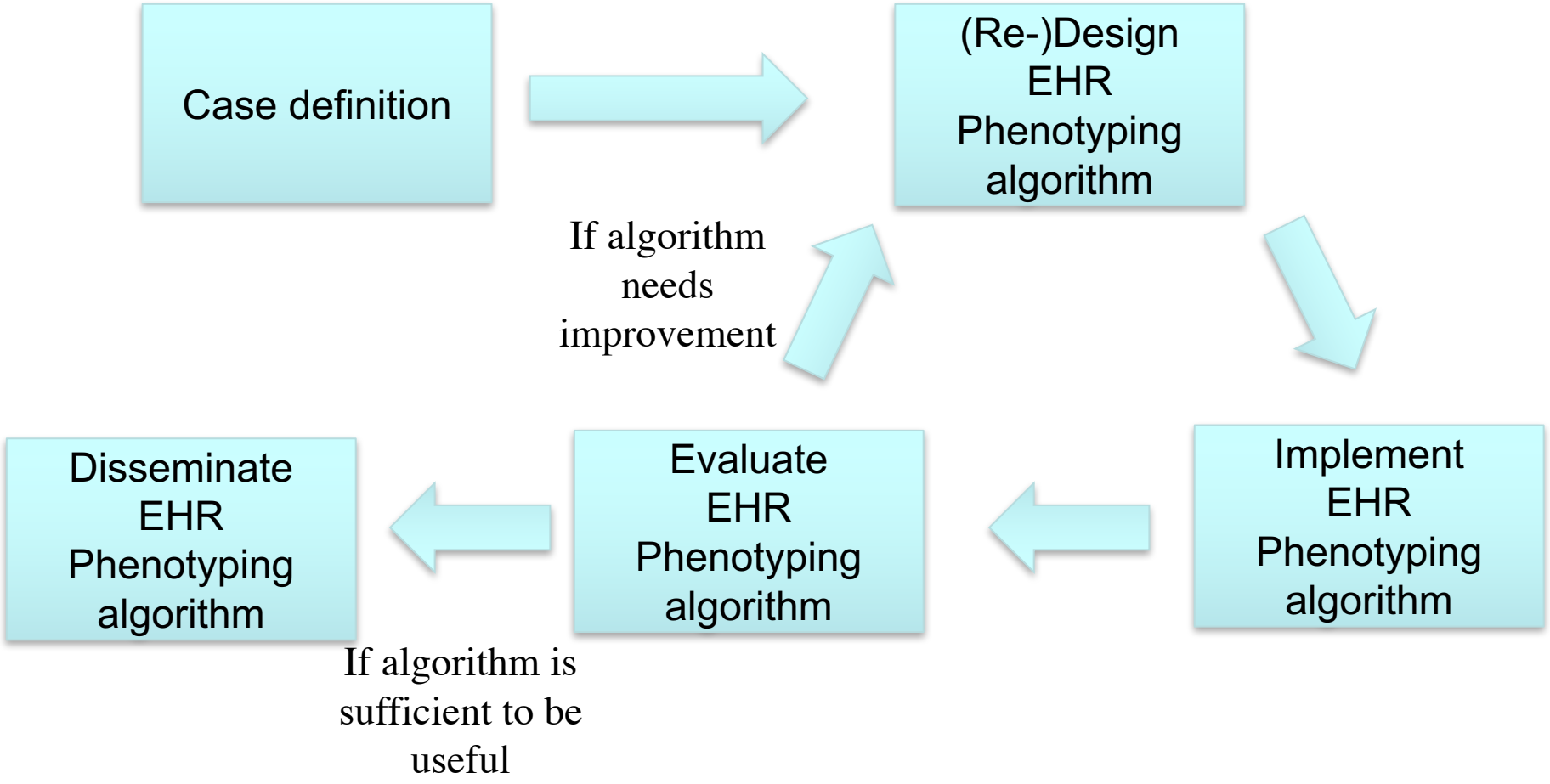
Overview of EHR phenotyping process



Overview of EHR phenotyping process

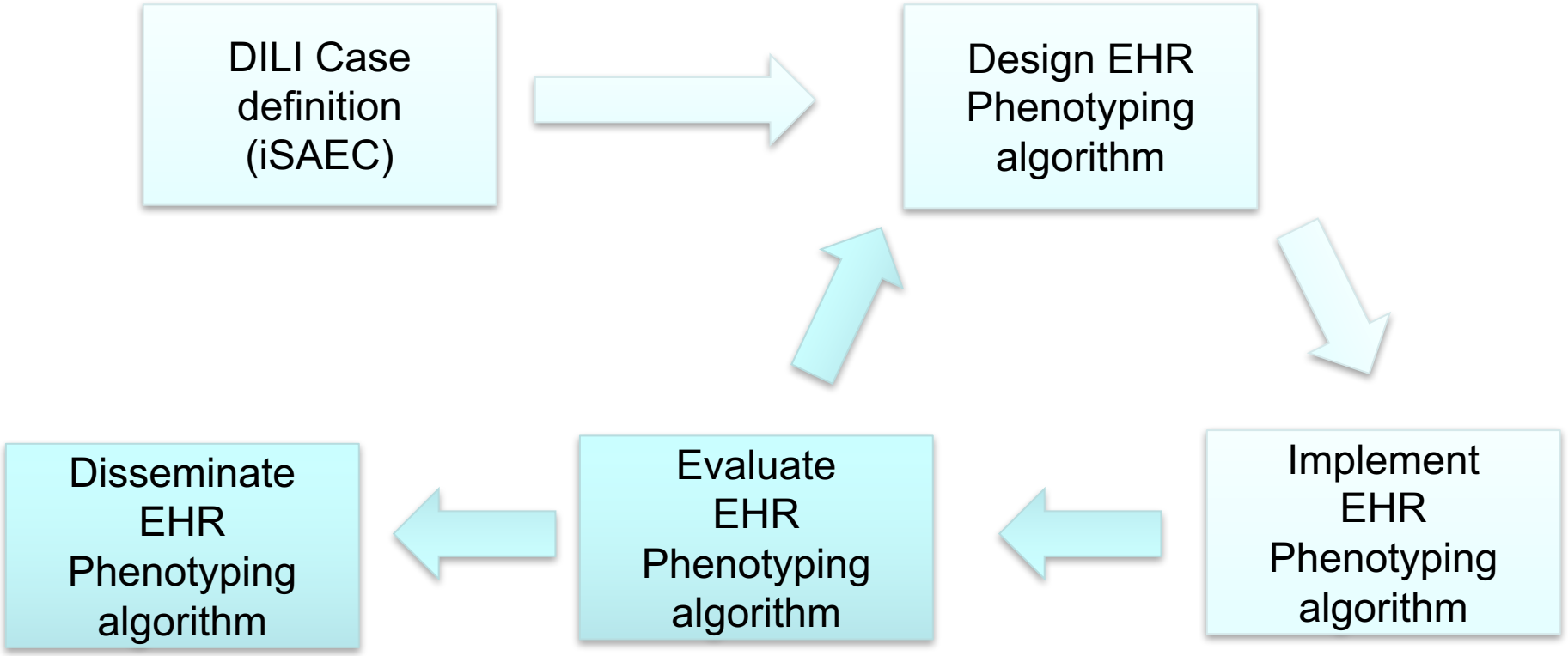


Overview of EHR phenotyping process



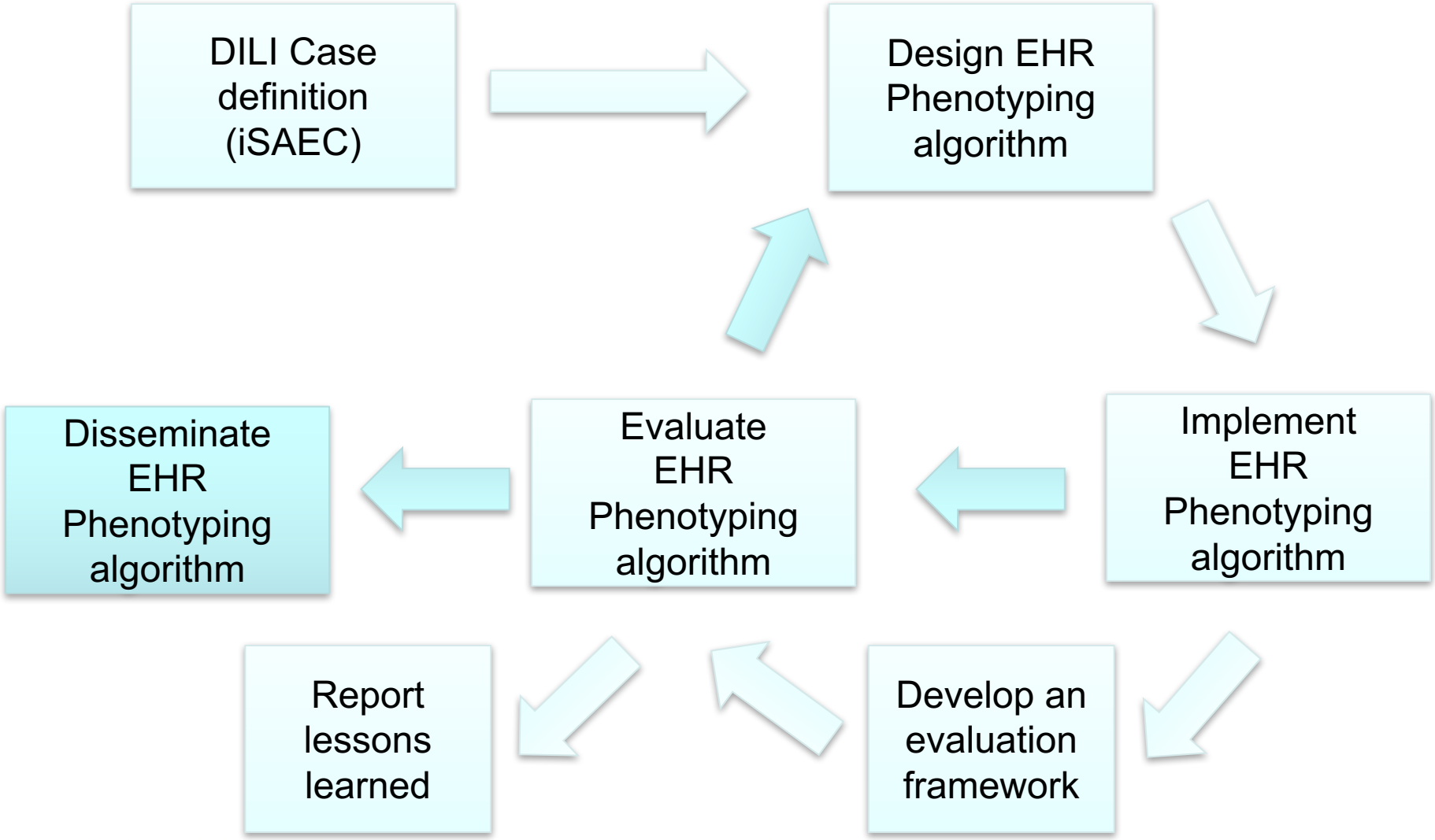
Overview of methods to develop & evaluate initial algorithm

JOHNS HOPKINS MEDICINE
emerge network
ELECTRONIC MEDICAL RECORDS AND GENOMICS

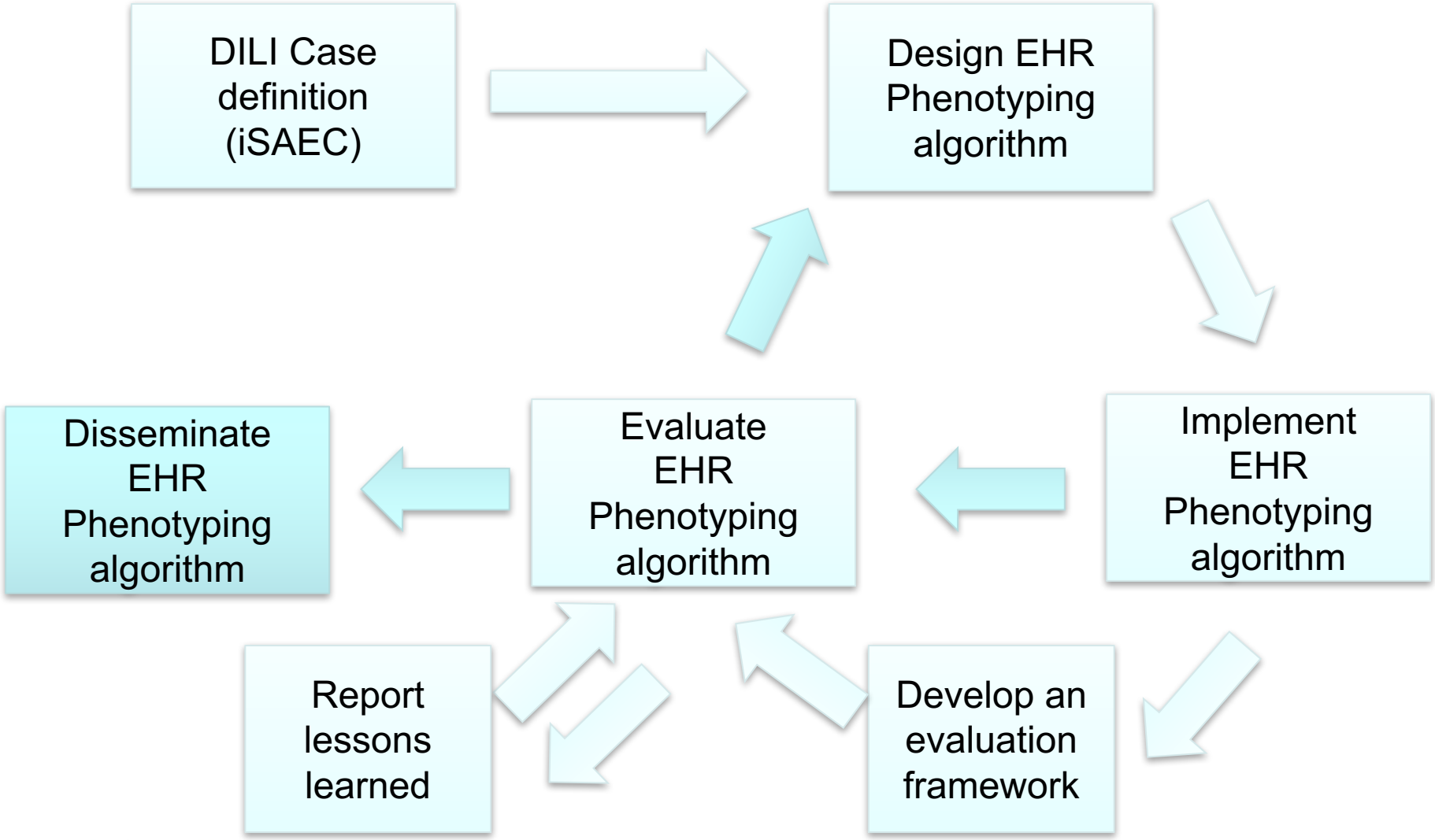


Overby, C. L., Weng, C., Haerian, K., Perotte, A., Friedman, C., & Hripcsak, G. (2013). Evaluation considerations for EHR-based phenotyping algorithms: a case study for drug-induced liver injury. *AMIA Summits on Translational Science Proceedings*, *2013*, 130.

Overview of methods to develop & evaluate initial algorithm



Lessons inform evaluator approach and algorithm design changes

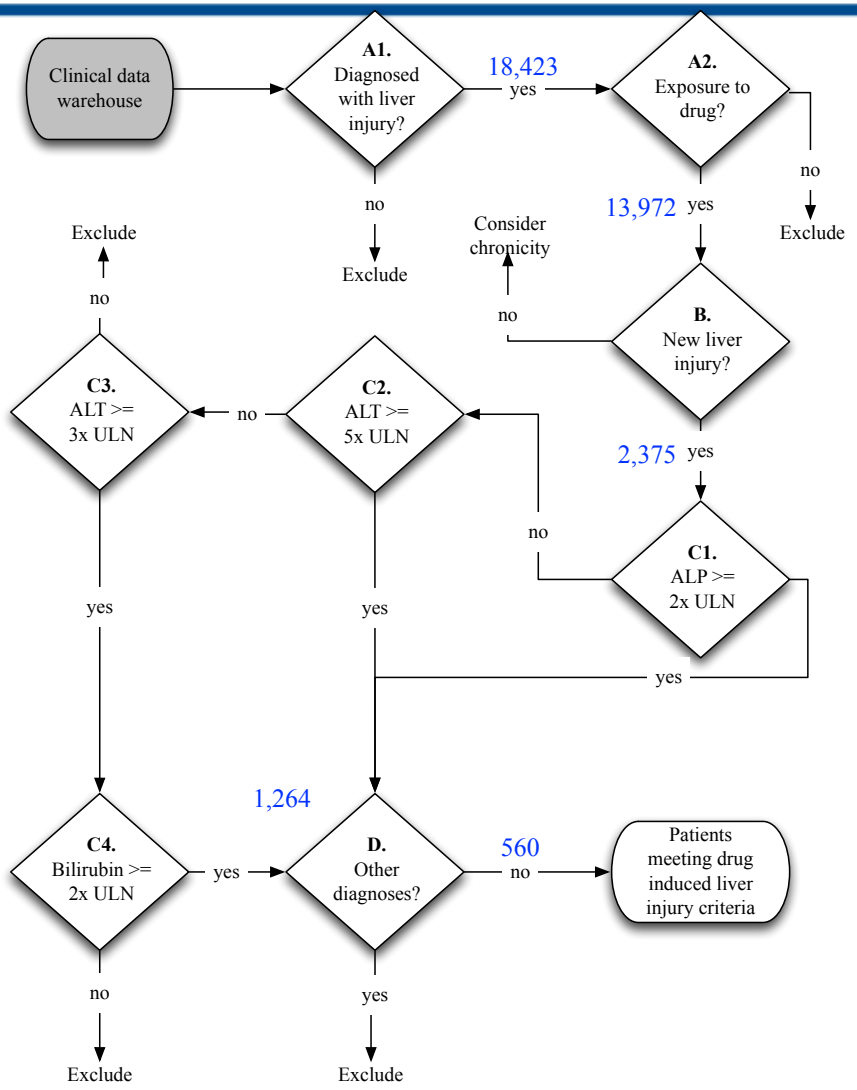


Initial DILI EHR phenotyping algorithm

DILI case definition

1. Liver injury diagnosis (A1)
   a. Acute liver injury (C1-C4)
   b. New liver injury (B)

2. Caused by a drug
   a. New drug (A2)
   b. Not by another disease (D)



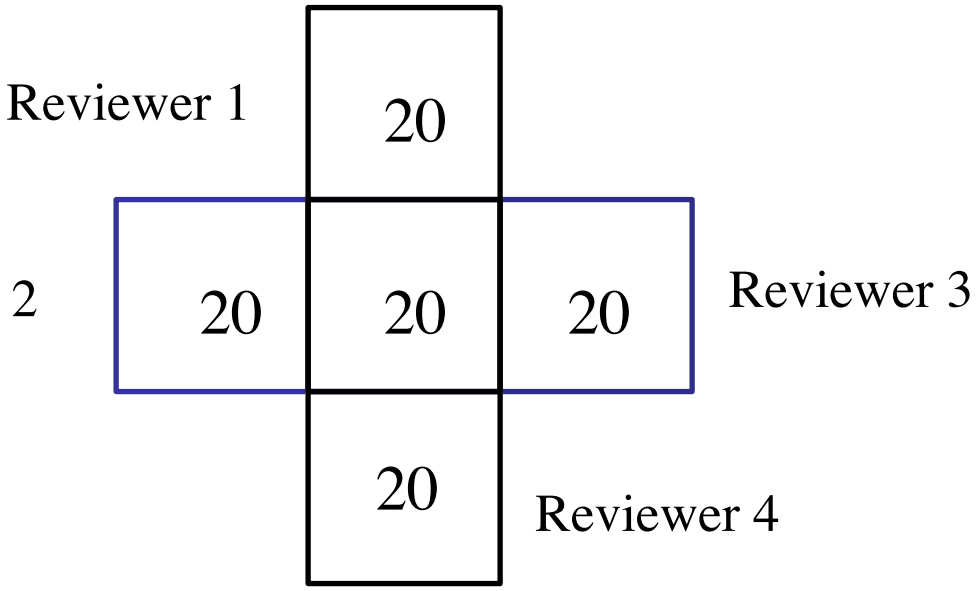
Ref: Aithal, G.P., et al. Case Definition and Phenotype Standardization in Drug-induced Liver Injury. Clin Charmacol Ther. 2011 Jun; 89(6):806-15

Estimated positive predictive value

Initial algorithm results:
100 randomly selected for manual review from 560 patients

- TP: 27
- FP: 42
- NA: 30

- PPV: TP/(TP+FP) = 27/(42+27) = **39%**



- Preliminary kappa coefficient: **0.50** (Moderate agreement)

- **Interpretation of PPV is unclear given moderate agreement among reviewers**

An evaluation framework and results

	Measurement study (evaluator effectiveness)	Demonstration study (algorithm performance)
Quantitative results	Kappa coefficient: 0.50	TP: 27 FP: 42 NA: 30 PPV: TP/(TP+FP) = 39%

Qualitative results	*Perceptions of evaluation approach effectiveness:*
	• Differences between evaluation platforms
	  • Visualizing lab values
	  • Availability of notes
	• Discharge summary vs. other notes
		Perceptions of benefit of results (themes in FPs):
		• Babies
		• Patients who died
		• Overdose patients
		• Patients who had a liver transplant

Capturing the value of data for use in clinical applications: some evaluation considerations

JOHNS HOPKINS MEDICINE

- Computational approaches that pull data from multiple sources is an iterative process (e.g. EHR phenotyping)
  - Complexity of the algorithm may influence
- Lessons learned from using an evaluation framework
  - What's correct for the algorithm may not be correct for the case definition (Are we measuring what we mean to measure?)
  - Evaluator effectiveness influences ability to draw appropriate inferences about algorithm performance
- Potential usefulness of an evaluation framework
  - Informs improvements in algorithm design
  - Informs improvements in evaluator approach
  - Likely more useful for rare and complex conditions

Characteristics of a test: What's important depends on context

- Sensitivity and specificity
- Electronic cohort
  - identification vs screening
- Monitor changes due to new healthcare practices or interventions
  - decision making vs inform policy

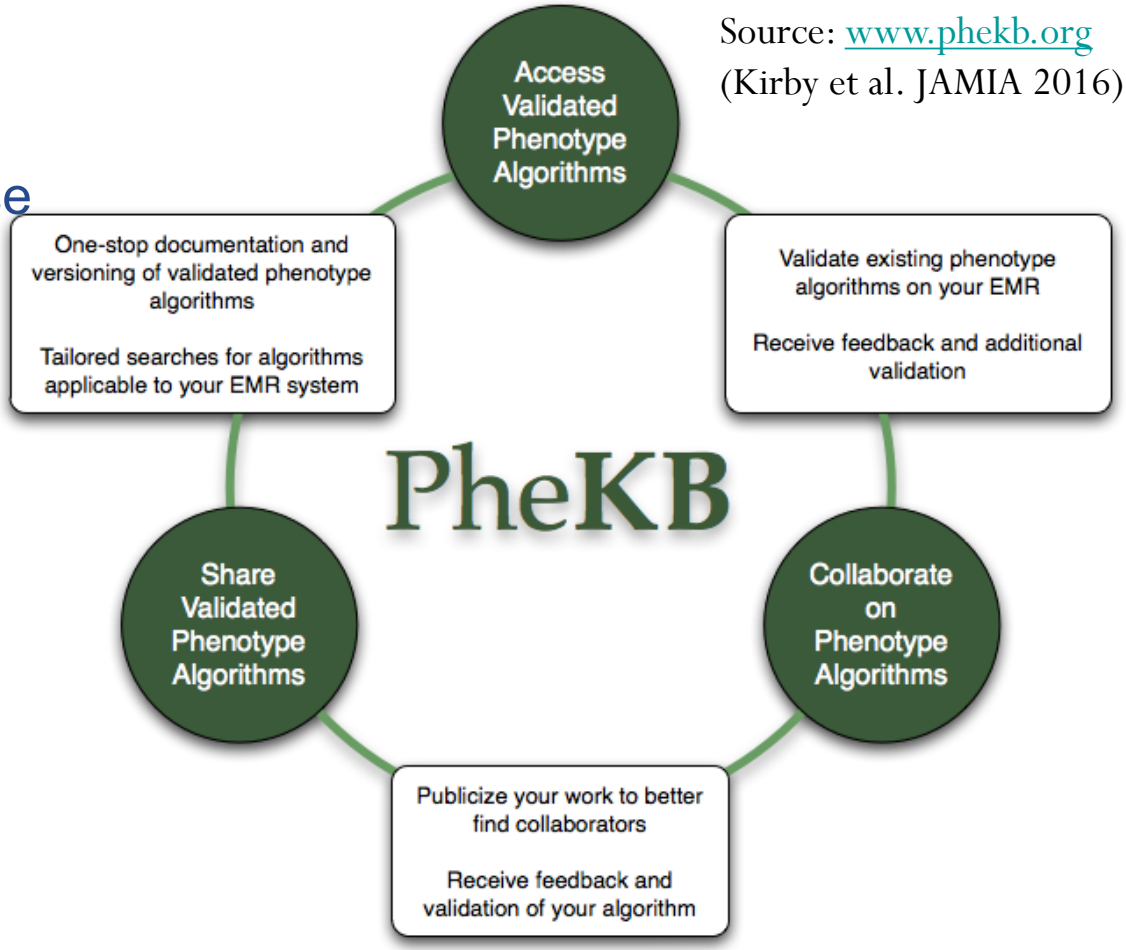
There have been many successes with extracting clinically relevant phenotypes from EHRs

JOHNS HOPKINS MEDICINE

- Type II diabetes
- Peripheral arterial disease
- Atrial fibrillation
- Crohn disease
- Multiple sclerosis
- Rheumatoid arthritis

- **High PPV**

Source: www.phekb.org
(Kirby et al. JAMIA 2016)



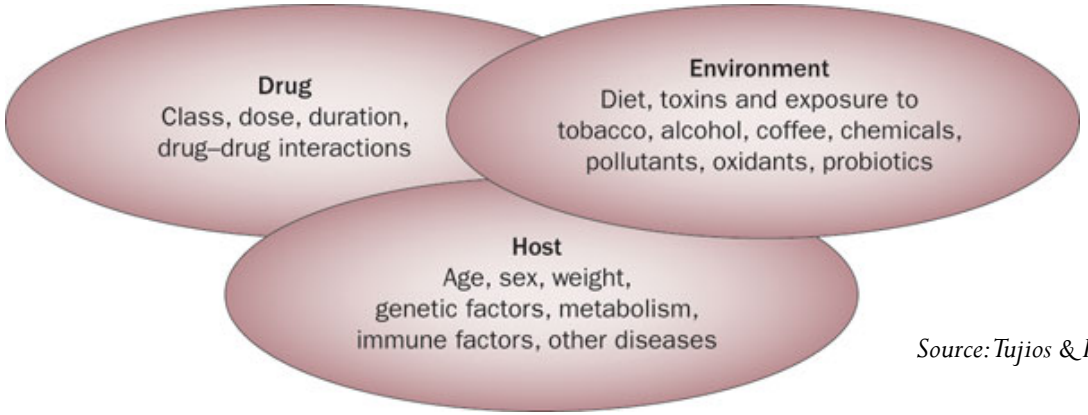
(Kho et al. JAMIA 2012; Kullo et al. JAMIA 2010; Ritchie et al. AJHG 2010; Denny et al. Circulation 2010; Peissig et al. JAMIA 2012)

emerge network
ELECTRONIC MEDICAL RECORDS AND GENOMICS

Data sharing can enable sample sizes needed for new discoveries

JOHNS HOPKINS MEDICINE

- Drug response is complex
  - Risk factors in pathogenesis of drug induced liver injury (DILI)



Source: Tujios & Fontana et al. Nat. Rev. Gastroenterol. Hepatol. 2011

- Sample sizes are small compared to typical association studies of common disease
  - Small prevalence
  - Several responder types

International SAE Consortium

DILIN Drug-Induced Liver Injury Network

emerge network
ELECTRONIC MEDICAL RECORDS AND GENOMICS

- Data sharing to achieve sample sizes needed for discovery

Challenges to leveraging current innovations in using big data to improve health systems

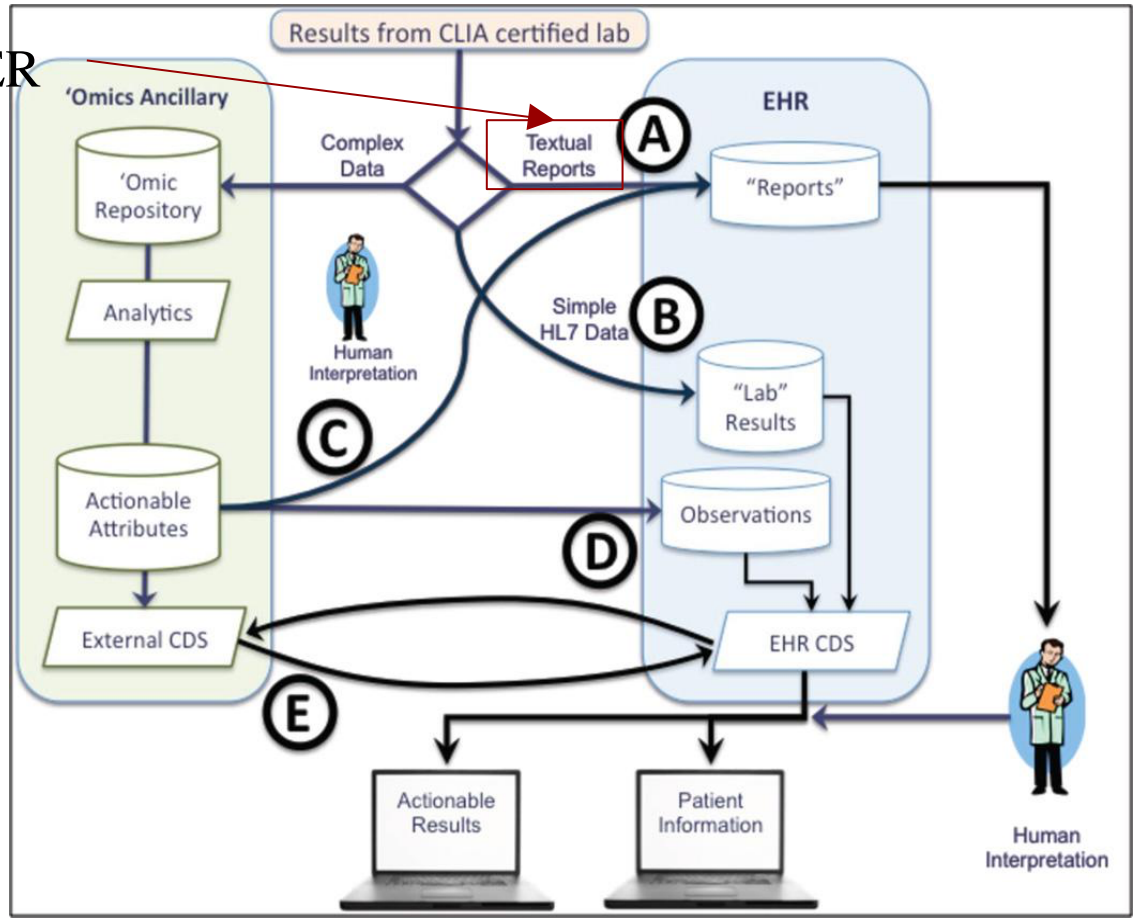
JOHNS HOPKINS MEDICINE

- Challenge #1: How can we decipher the meaning of data collected from various sources?

- Challenge #2: How can we deliver new evidence from (big) data analyses in an effective way?

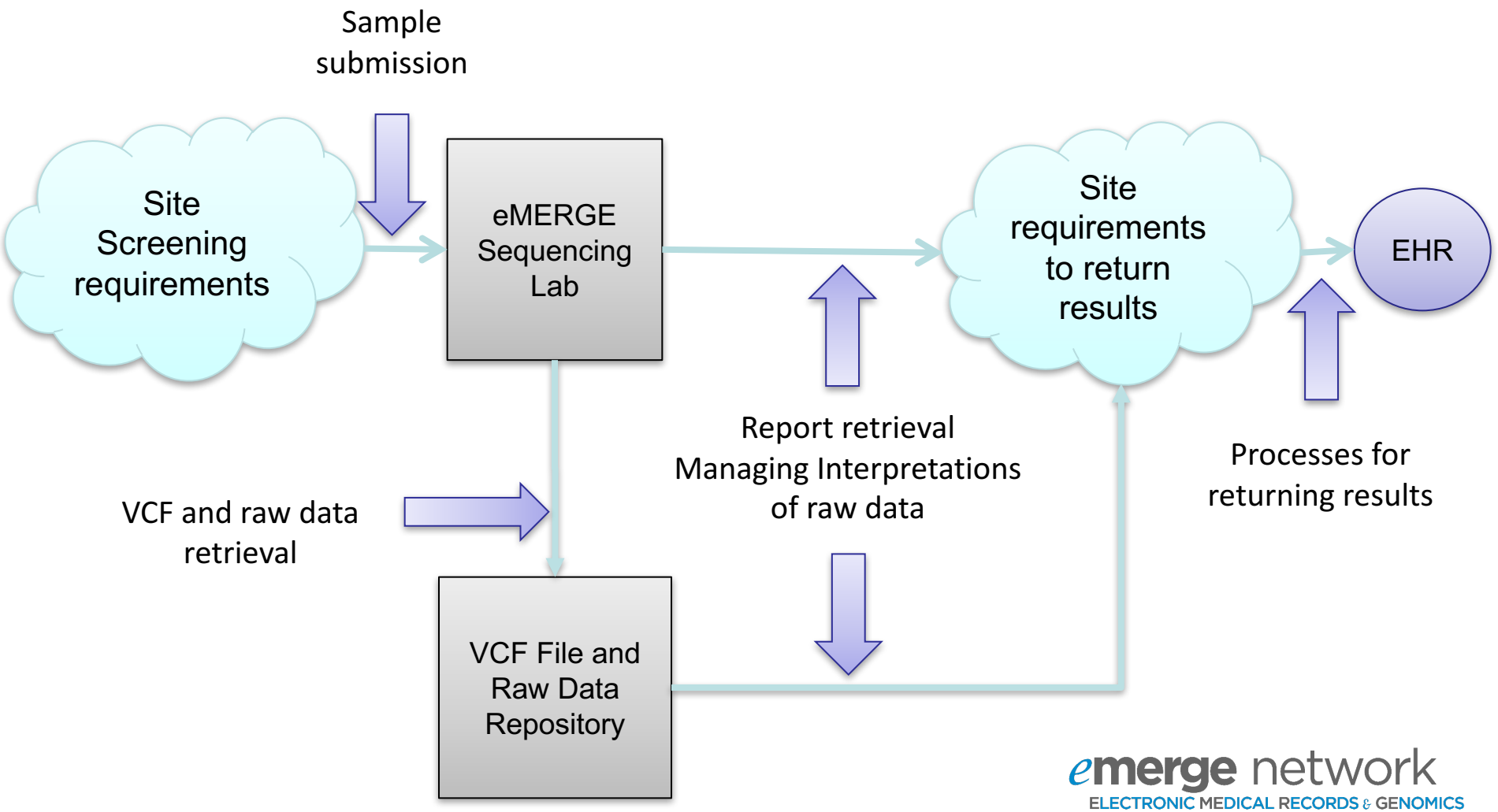
Previous work outlining data pathways to deliver genetic and genomic test results to healthcare providers

eMERGE II & CSER
(Shirts et al. 2015)



(Starren, Williams, Bottinger, JAMA. 2013)

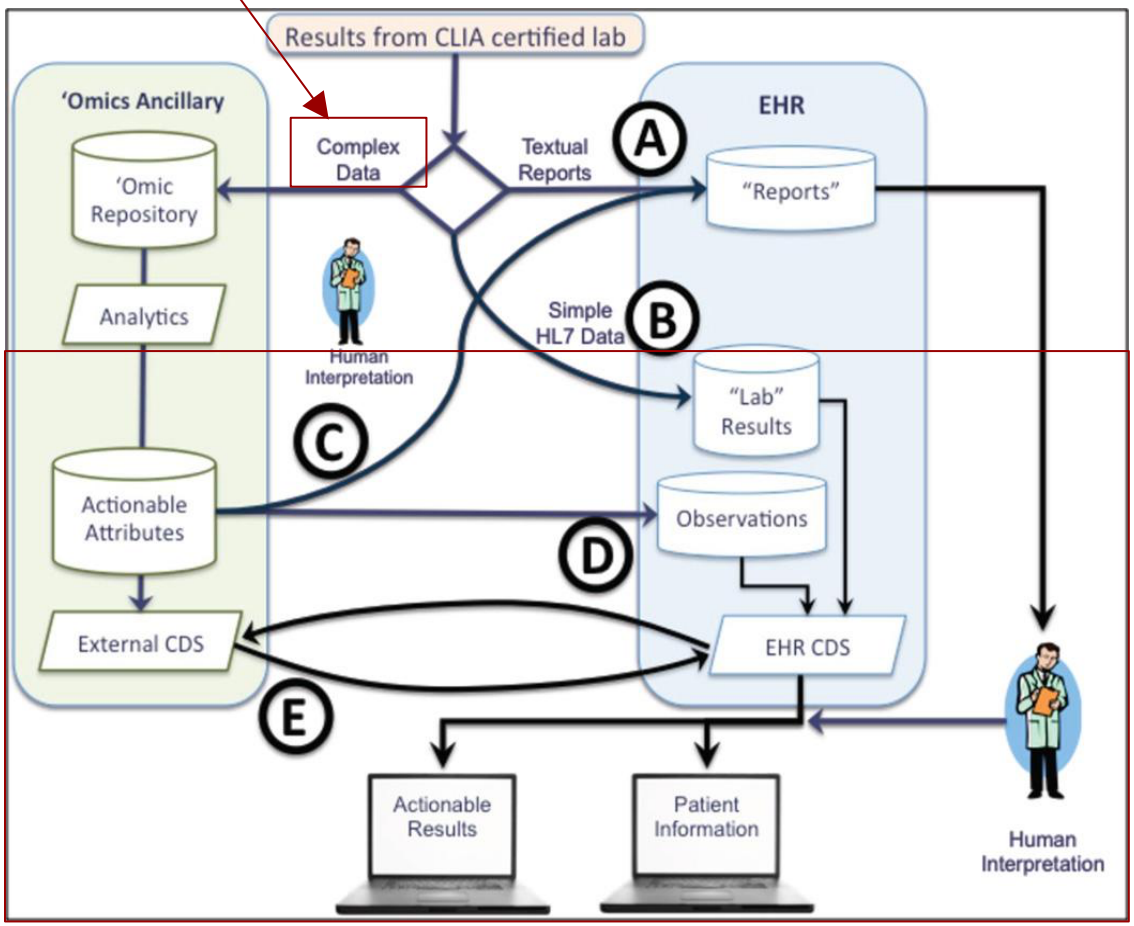
Enabling complex, structured genetic and genomic test results to be returned



Shift to receiving genetic and genomic test results without previous knowledge of how to interpret clinically

JOHNS HOPKINS MEDICINE

eMERGE III



Challenges illustrating a need for replicable and reproducible data interpretation

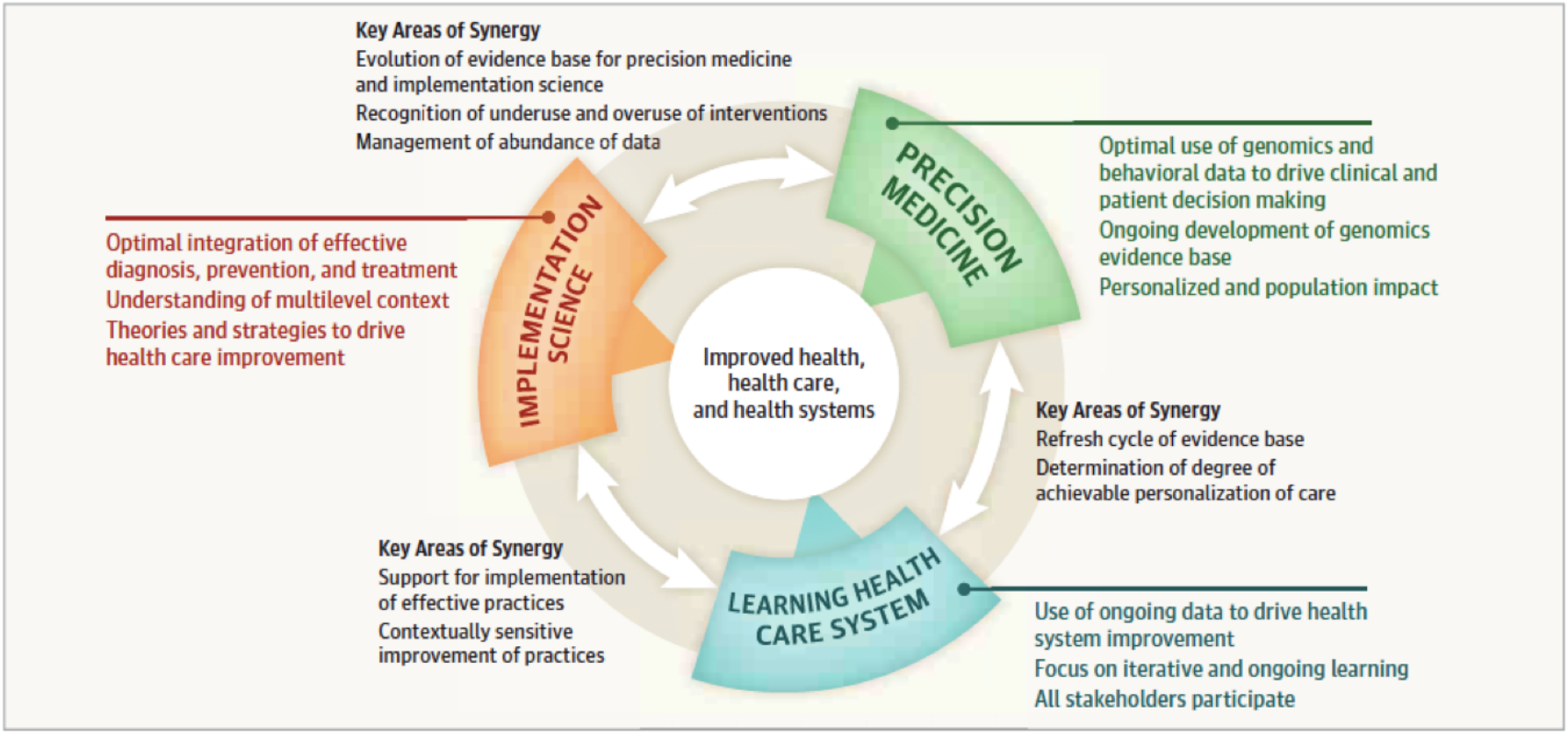
JOHNS HOPKINS MEDICINE

- Replicability
  - Genomic variant interpretations may change
  - Clinical guidelines may change
- Reproducibility
  - Use of calculations at multiple institutions

Research and practice co-exist to enable ongoing learning and evidence development – replicability & reproducibility are important

JOHNS HOPKINS MEDICINE

Figure. Contributions of Implementation Science, Learning Health Care System, and Precision Medicine



(Chambers, Feero, Khoury, JAMA. 2016)

Example: Reclassification of variant interpretation over time

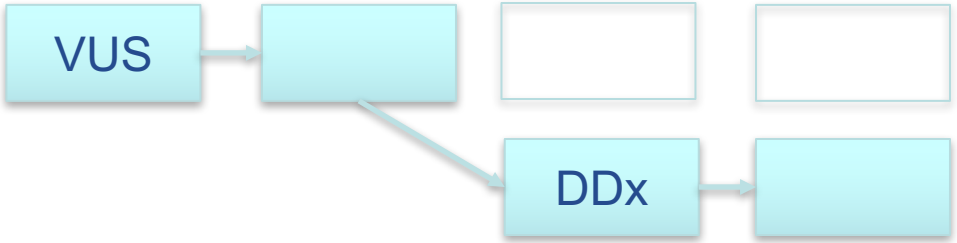
- **VUS Definition:** Genetic variants that cannot be classified definitively as pathogenic or benign at this time. Many are missense sequence variants that alter a single amino acid or in noncoding portions of genes. Many VUS are previously undescribed novel variants. VUS are reported on a variety of genetic testing platforms. Over time, VUS may be reclassified as benign or pathogenic; however, laboratories differ in whether VUS results are amended on clinical reports.

- **Clinical use example:** A 43-year-old female patient with a personal and family history of breast cancer undergoes sequencing analysis of BRCA1 and BRCA2. A missense VUS is reported in BRCA1 and reported as a VUS. Therefore **it is not recommended that testing for this variant be used to determine risk in relatives of this patient. Nine months later, a revised laboratory report reclassifies the variant as pathogenic based on additional evidence. The EHR is updated to now follow the recommendations found in Diagnostic and Actionable categories.**

(Shirts et al. JAMIA 2015, Table 1)

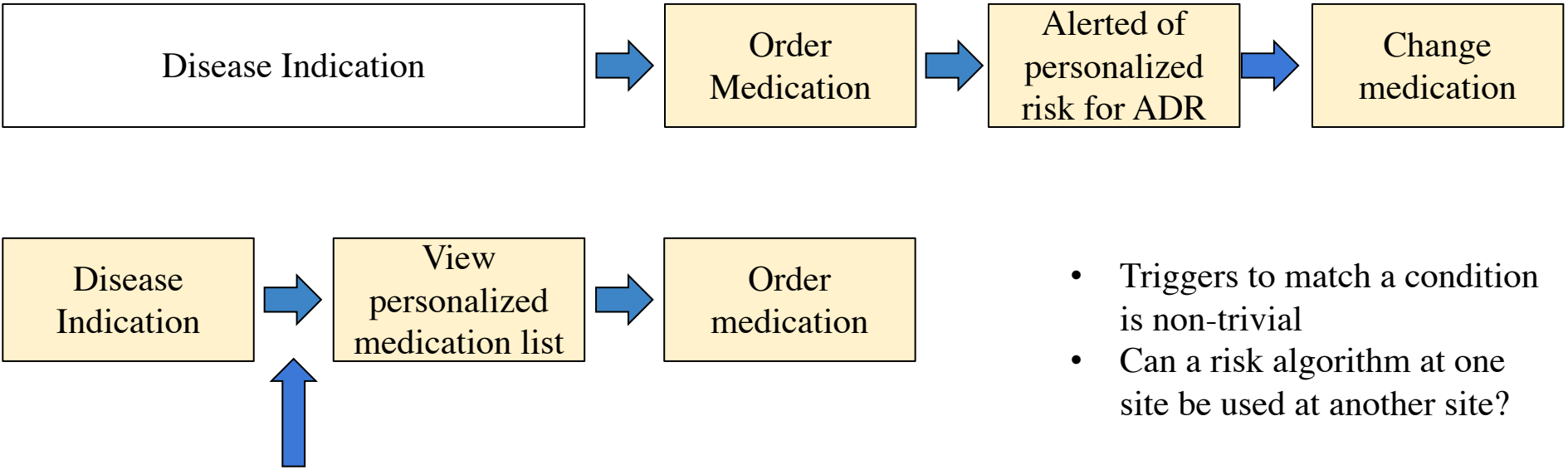
Implications for a learning healthcare system & importance of replicability

- What changes have occurred?
- When were changes made?
- How do changes influence retrospective data analyses?
- What is the impact of changes?



- **Tools to track provenance are needed**

Future use case: Upstream patient risk screening to inform prescribing decisions



- Triggers to match a condition is non-trivial
- Can a risk algorithm at one site be used at another site?

Implications for a learning healthcare system & importance of reproducibility

JOHNS HOPKINS MEDICINE

- Biggest challenge is data access (common challenge for clinical datasets)
  - Required to test reproducibility

- Potential solutions
  - Environment to assess models with different data
  - New data governance models (e.g., Sage Bionetworks, *Wilbanks & Friend, Nat Biot, 2016*)
  - Synthetic datasets (e.g., C. Chute, NCATS “Translator” grant)

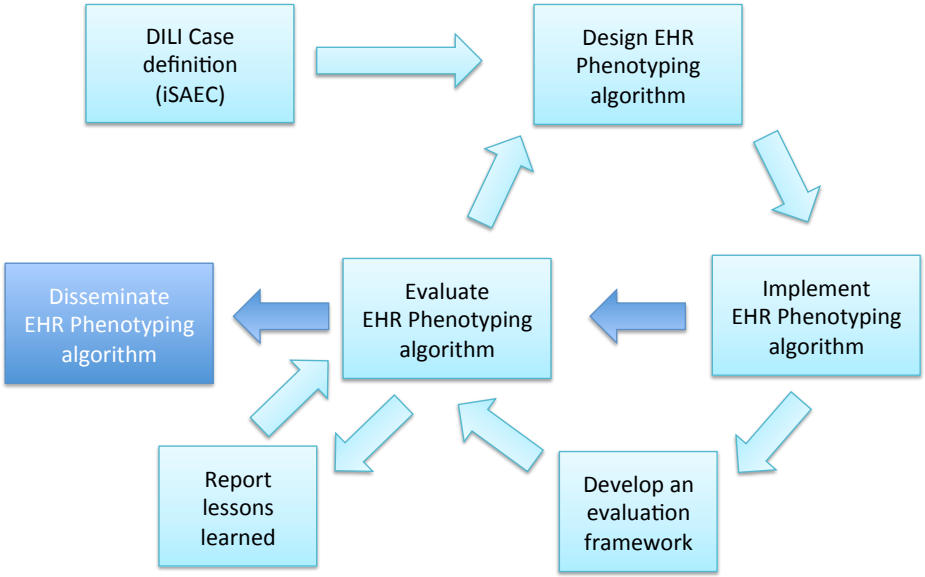
Summary of points

- **Challenge #1: How can we decipher the meaning of data collected from various sources?**
- Computational approaches that pull data from multiple sources is an iterative process (e.g. EHR phenotyping)
  - Complexity of algorithm may influence
  - Context influences value
  - Evaluation approach & threshold depends on context
- **Challenge #2: How can we deliver new evidence from (big) data analyses in an effective way?**
- Replicability of data interpretations are needed to enable a learning healthcare system
  - Re-interpretation of test results is a new paradigm and thus current healthcare systems are not designed to capture change over time
  - Capturing provenance may help
- Reproducibility of findings are needed to validate big data applications
  - Data access is a major challenge
  - Analytic environment, planning for data sharing and use of synthetic data may help

Acknowledgments

- Columbia mentors (Chunhua Weng , George Hripcsak, Yufeng Shen)
- Columbia collaborators (Carol Friedman, Krystl Haerian, Adler Perotte)
- Sandy Aronson (eMERGE Phase III EHRI WG co-Chair)
- eMERGE Phase III EHRI WG members
- Brian Shirts (University of Washington, CSER EHR WG member)
- Christopher Chute (Johns Hopkins University, NCATS Translator TransMed co-PI)

- Funding
  - Columbia Training in Biomedical Informatics (NIH NLM #T15 LM007079)
  - eMERGE Phase II: U01 HG006380 (Mount Sinai), HG006379 (Mayo Clinic),
  - eMERGE Phase III: U01HG8679 (Geisinger Health Clinic)
  - NIH NCATS Translator TransMed project grant (NIH NCATS 1OT3 TR00201)



JOHNS HOPKINS MEDICINE

Casey L. Overby
overby@jhu.edu

Challenge #1: How can we decipher the meaning of data collected from various sources?

Challenge #2: How can we deliver new evidence from (big) data analyses in an effective way?