

# Getting the Most out of RWD through Enrichment of the Data

Cecil O. Lynch, MD, MS

Global Biomedical Informatics Lead

Accenture, LLP



## Agenda

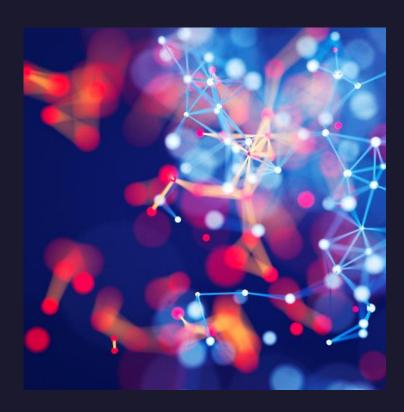
RWD Definition for this talk

Limitations of RWD

**Enrichment Definition** 

Examples

The Future









#### Introduction

• RWD from EMR and Claims is driving significant value in Life Sciences to identify patient journeys, adverse events, market penetration patterns and potential identification of clinical trial candidates among other uses. This talk will focus on getting even more value out of RWD by enriching that data.

## What do we mean by Enrichment?

Extending data through linkage to other data sets where there are attributes that do not exist in the original RWD.



Can we get more out of the data through enrichment?

Let's see a few examples.



### SDOH Example

#### **ADVANTAGES**

- Be able to adjust the messaging, therapies and interventions to meet the needs of the individual.
- Engage underserved populations
- Be able to potentially broaden the clinical trial participant pool.
- HL7 FHIR Accelerators defining attributes

#### **ISSUES**

- Need to have detailed zip code or geo-location data to connect the data
- Wide variance in attributes to consider (food availability tailored to the individual, financial data, health care services, cognitive ability),

### Genetics Example

#### **GENETIC TO DISEASE CONNECTION**

Hypertrophic cardiomyopathy  Essential hypertension Risk Factor 31.558 67.09% 8,868,477 Hyperlipidemia 24,619 52.34% 7,466,066 Congestive heart failure Pathogenic 13,612 28.94% 1,205,105 Cardiomegaly Pathogenic 12,864 27.35% 644,360 Atrial fibrillation Pathogenic 12,514 26.60% 1,462,021 Cardiac arrhythmia Pathogenic 12,295 26.14% 3,336,759 Anemia Pathogenic 12,295 26.14% 3,336,759 Gastroesophageal reflux disease Pathogenic 11,891 25.28% 4,275,874 Type 2 diabetes mellitus without complication 11,754 24.99% 3,229,169 Atherosclerosis of coronary artery without angina pectoris Gastroesophageal reflux disease without esophagitis 11,551 24.56% 2,858,011 Primary cardiomyopathy  Fields to Display Gene Disease Associations  Condition Term Hypertrophic cardiomyopathy Score	Population Comorbidity Analysis														
Hyperlipidemia	Condition Term	Comorbidity Terms				No. Comorbid		id %	Comorbid			% Fre	quency erse)		
Congestive heart failure	Hypertrophic	Essential hypertension			Risk Factor				58	67.09%	8,868,477			15.679	
Cardiomegaly	cardiomyopathy	Hyperlipidemia						24,6	19	52.34%	7	,466,066	5	10.819	
Atrial fibrillation		Congestive heart failure			Pathogenic			13,6	12	28.94%	1	,205,105	5	2.159	
Cardiac arrhythmia		Cardiomegaly			Pathogenic		12,864		54	27.35%	644,3		)	1.189	
Anemia Pathogenic 12,005 25.52% 3,777,923   Gastroesophageal reflux disease Pathogenic 11,891 25.28% 4,275,874   Type 2 diabetes mellitus without complication		Atrial fibrillation			Pathogenic		12,514		14	26.60%	1,462,02		ř.	2.519	
Gastroesophageal reflux disease   Pathogenic   11,891   25.28%   4,275,874		Cardiac arrhythmia			Pathogenic			12,2	95	26.14%	3,336,759		)	2.419	
Type 2 diabetes mellitus without complication Atherosclerosis of coronary artery without angina pectoris Gastroesophageal reflux disease without esophagitis Primary cardinmyonarhy  Gene Disease Association Variant Disease Association Gene Name Gene Variant Pathogenic  MYH7 Pathogenic  11,754 24.99% 3,229,169 11,632 24.73% 1,424,730 11,632 24.73% 1,424,730 24.56% 2,858,011 24.56% 24.56% 24.56% 24.56% 24.56% 2.858,011 24.56% 28.58,011 24.56% 24.5		Anemia			Pathogenic			12,0	05	25.52%	3,777,923		3	4.679	
Atherosclerosis of coronary artery without angina pectoris  Gastroesophageal reflux disease without esophagitis  Primary cardiomyonathy  Gene Disease Association  Variant Disease Association  Gene Variant  Gene Variant  Pathogenic  MYH7  Pathogenic  11,632  24.73%  1,424,730  24.56%  2,858,011  10.402  22.11%  335.250  Condition Term  Hypertrophic cardiomyopathy  Score  Links  Score  Links  Score  Links  Score  Links  Score  1 0.87  rs1060501452  1 0.87  rs121913624  1 0.87  1 0.87  1 0.87  rs121913624  1 0.87  1 0.87		Gastroesophageal reflux disease			Pathogenic			11,8	91	25.28%	4,275,874			6.119	
pectoris Gastroesophageal reflux disease without esophagitis Primary cardiomyonathy  Gene Disease Association Variant Disease Association Gene Name Gene Variant Pathogenic  MYH7 Pathogenic  Description  MYH7 Pathogenic  MYH7 Pa		Type 2 diabetes mellitus without complication						11,7	54	24.99%	3,229,169		)	5.909	
Primary cardinmyonarhy    Tourise   Primary cardinmyonarhy   Tourise   Primary cardinmyonarhy   Tourise   Primary cardinmyonarhy   Tourise   Primary cardinmyonarhy   Prima								11,6	32	24.73%	1,424,730			2.609	
Fields to Display  Gene Disease Association  Variant Disease Association  Gene Disease Associati		Gastroesophageal reflux disease without esophagitis						11,551		24.56%	2,858,011			5.229	
Gene Disease Association Variant Disease Association Gene Name Pathogenic  MYH7  Pathogenic  Condition Term Hypertrophic cardiomyopathy Score Links Score Links Score Links Score Links Score 1 0.87 1 0.87 1 0.87 1 0.87 1 0.87 1 0.87 1 0.87 1 0.87 1 0.87 1 0.87 1 0.87 1 0.87 1 0.87		Primary cardiomyonathy			10.402		12	22.11%	335.250			0.619			
Variant Disease Association Gene Disease Association Gene Name Gene Variant Pathogenic  MYH7 Pathogenic  Pathogenic  MYH7 Pathogenic  Pathogenic  Pathogenic  MYH7 Pathogenic  MYH7 Pathogenic  MYH7 Pathogenic  MYH7 Pathogenic  MYH7 Pathogenic  MYH7  Pathogenic  MYH	Fields to Display	Gene Disease Associations													
Variant Disease Association   Gene Disease Association   Gene Disease Association   Gene Name   Variant Disease Association   Gene Variant   Links   Score   Links   Score   Core   Links   Score   Links   Links   Score   Links	Gene Disease Association	-				Con	Condition Term		Hype	-lypertrophic cardiomyopathy					
□ Pathogenic □ MYH7 □ Pathogenic □ msv513807 1 0.8	The same and the same	Gene Disease Association	Gene Name	Variant Di	t Disease Association						, -  ,	-	Score		
rs1060501452 1 0.87 1 0.87 rs121913624 1 0.87 1 0.87	_	□ Pathogenic □ MYH7 □ Patho			enic	nsv	nsv513807		1		0.87	1	0.87		
rs121913624 1 0.87 1 0.87		- autogette					rs1060501452		1	-					
1312 13 13023									1						
rs121913626 1 0.87 1 0.87									1				0.87		
									1			1	0.87		

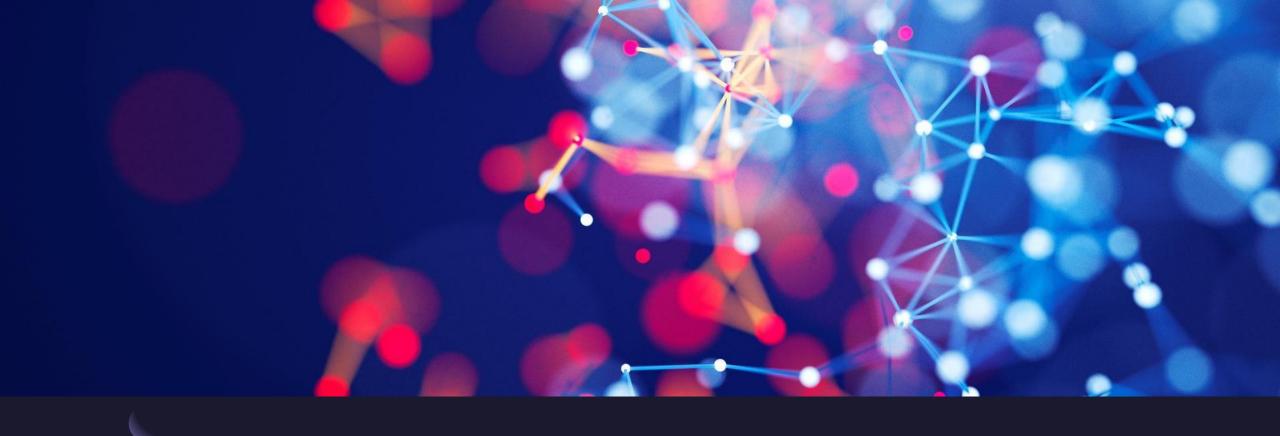
#### **DETAILS**

- Many sources of data available to link a condition to a gene
- Need to standardize the condition data across a large variance in representations (ICD10, SNOMED, others),
- Once the gene is linked can extend to variants, proteins, biopathways

## Where does the value of genetic enrichment of Big Data come in?

- Identifying novel genetic relationships between conditions
- Estimating the prevalence of mutations
- Identifying disease occurrence of rare disease based on profiles of built on genetically related comorbidities that pre-date the disease od interest,

- Improving tertiary genetic analysis by identifying suspect uncertain variants that may well be pathogenic based on frequency of co-occurring comorbid conditions
- Identifying new targets for therapies based on related biopathways identified through comorbid genetic relations that point to new etiologies



## Summary

The enrichment of RWD has many benefits that go beyond the initial data use cases and is economically justified.

Gen Al is going to open up many other benefits of the enrichment of RWD.

#### Thank You

Cecil O, Lynch, MD, MS

cecil.o.lynch@accenture.com

