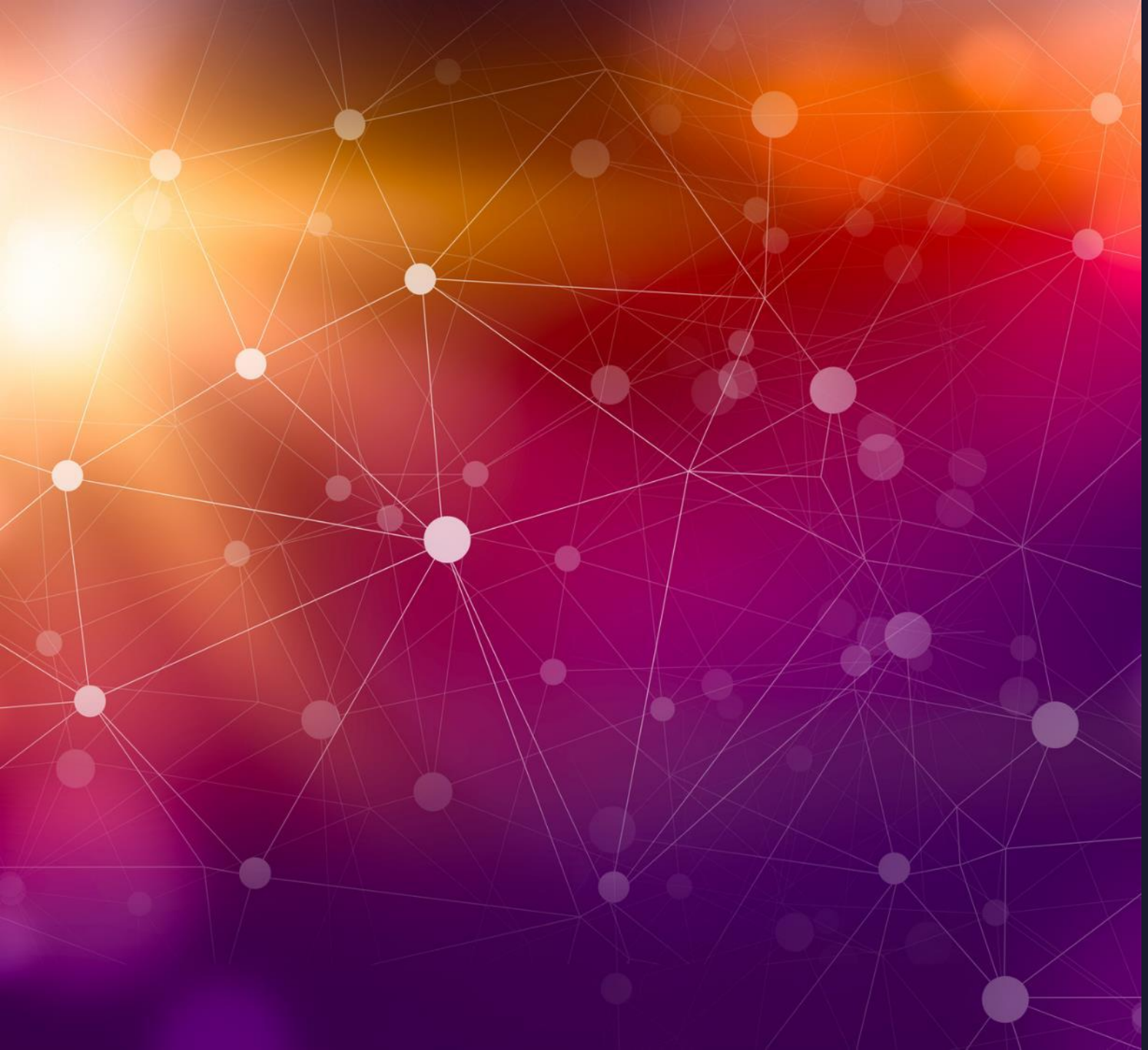


# Getting the Most out of RWD through Enrichment of the Data

Cecil O. Lynch, MD, MS

Global Biomedical Informatics Lead

Accenture, LLP

# Agenda

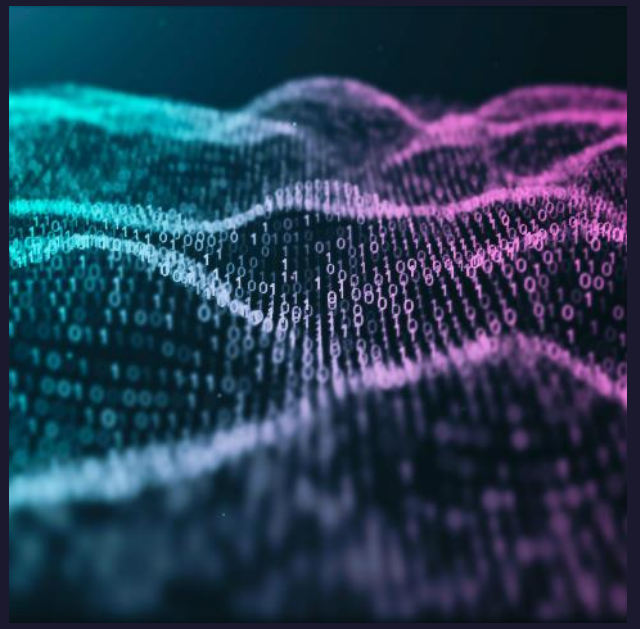
RWD Definition for this talk

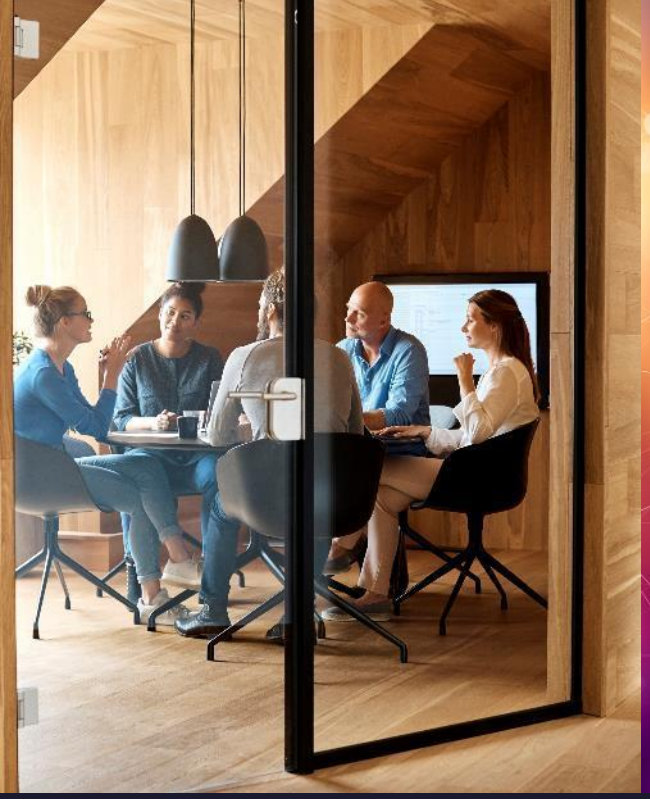
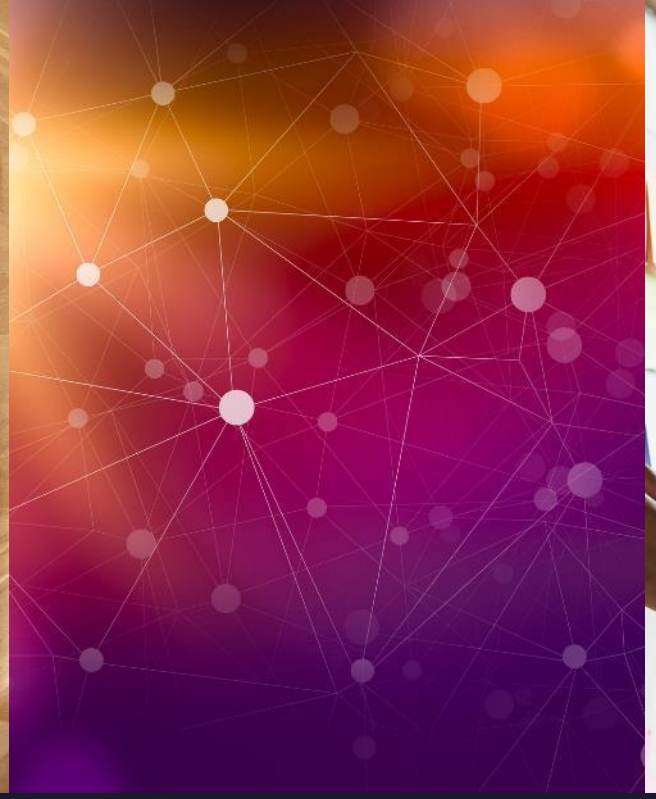
Limitations of RWD

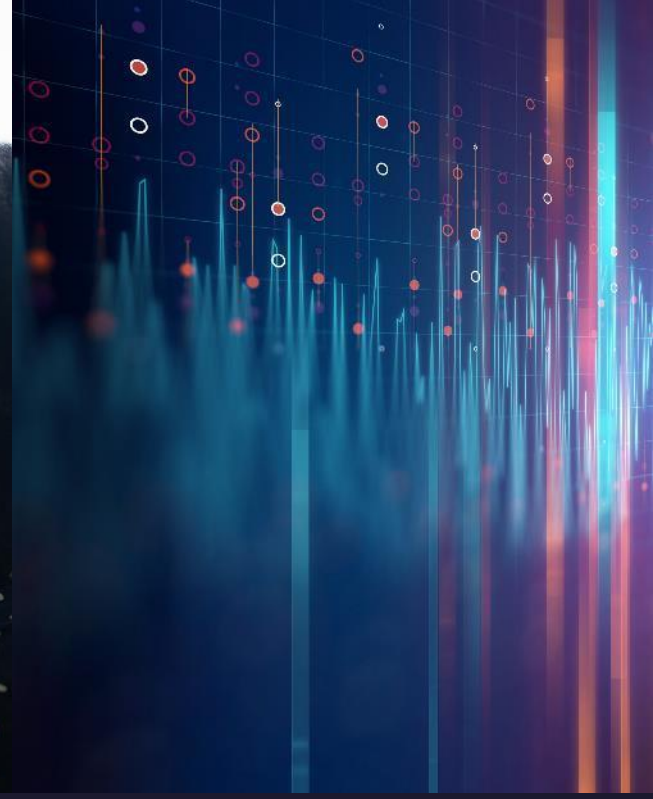
Enrichment Definition

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

**Population Comorbidity Analysis**

| Condition Term | Comorbidity Terms | Variant Disease Association | No. Comorbid | % Comorbid | No. Patients Universe | % Frequency (Universe) |
|---|---|---|---|---|---|---|
| Hypertrophic cardiomyopathy | Essential hypertension | Risk Factor | 31,558 | 67.09% | 8,868,477 | 15.67% |
| | Hyperlipidemia | | 24,619 | 52.34% | 7,466,066 | 10.81% |
| | Congestive heart failure | Pathogenic | 13,612 | 28.94% | 1,205,105 | 2.15% |
| | Cardiomegaly | Pathogenic | 12,864 | 27.35% | 644,360 | 1.18% |
| | Atrial fibrillation | Pathogenic | 12,514 | 26.60% | 1,462,021 | 2.51% |
| | Cardiac arrhythmia | Pathogenic | 12,295 | 26.14% | 3,336,759 | 2.41% |
| | Anemia | Pathogenic | 12,005 | 25.52% | 3,777,923 | 4.67% |
| | Gastroesophageal reflux disease | Pathogenic | 11,891 | 25.28% | 4,275,874 | 6.11% |
| | Type 2 diabetes mellitus without complication | | 11,754 | 24.99% | 3,229,169 | 5.90% |
| | Atherosclerosis of coronary artery without angina pectoris | | 11,632 | 24.73% | 1,424,730 | 2.60% |
| | Gastroesophageal reflux disease without esophagitis | | 11,551 | 24.56% | 2,858,011 | 5.22% |
| | Primary cardiomyopathy | | 10,402 | 22.11% | 335,250 | 0.61% |

Fields to Display

- Gene Disease Association
- Variant Disease Association
- Gene Name
- Gene Variant

**Gene Disease Associations**

| | | | Condition Term | Hypertrophic cardiomyopathy | | Score | |
|---|---|---|---|---|---|---|---|
| Gene Disease Association | Gene Name | Variant Disease Association | Gene Variant | Links | Score | Links | Score |
| Pathogenic | MYH7 | Pathogenic | nsv513807 | 1 | 0.87 | 1 | 0.87 |
| | | | rs1060501452 | 1 | 0.87 | 1 | 0.87 |
| | | | rs121913624 | 1 | 0.87 | 1 | 0.87 |
| | | | rs121913625 | 1 | 0.87 | 1 | 0.87 |
| | | | rs121913626 | 1 | 0.87 | 1 | 0.87 |
| | | | rs121913627 | 1 | 0.87 | 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 AI 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