

Use of FHIR in the Generation of Real World Evidence



Edinburgh 2017

Paper Authors

Regina Zopf, FDA, Bethesda, MD, USA
Jeff Abolafia, Rho, Chapel Hill, NC, USA
Bhargava Reddy, UCB, Raleigh, NC USA

Presented by

Trisha D. Simpson



Fast Healthcare Interoperability Resources

What we did

- Extracted data for diabetic type 2 patients from the MITRE SyntheticMass Synthea Health Information Exchange (HIE)
- Mapped certain data points to CDASH & SDTM
- Generated CDASH-compliant CRFs
- Generated SDTM-compliant datasets
- **Had a little fun**
- Ran across a few needed tweaks to FHIR, CDASH & SDTM
- Realized further exploration & more work is needed

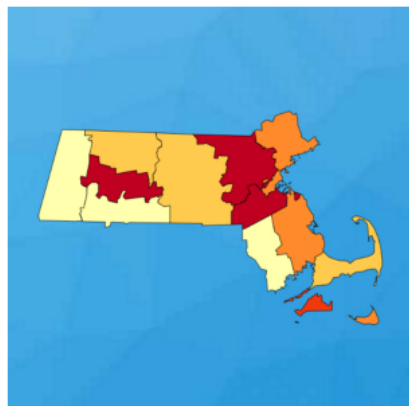
Why we did it

- EHR data are potential goldmines for both RWE & clinical research
- FHIR Resources are similar to the structure used in CRFs
- We wanted to see how difficult it is to populate CRFs & datasets
- **We wanted to have fun**
- FHIR shows promise as a way to link provider and research data
- Potential benefits include identification of related conditions/interactions and time reduction for data entry at the site
- If we *really* want to get to bi-directional data integration (Healthcare  Research) and true interoperability, **FHIR** is a solution

The real reason ...



SyntheticMass Synthea Database



About SyntheticMass

Background

[SyntheticMass](#) contains a million realistic but fictional residents of the state of Massachusetts. The synthetic population statistically mirrors the real population in terms of demographics, disease burden, vaccinations, medical visits, and social determinants. SyntheticMass establishes a risk-free environment for experimenting with:

- Data visualization
- Risk stratification
- Care management
- Clinical decision support
- Patient reported data integration
- Evaluation of new treatment models
- Privacy and consent models
- Security and authorization models
- Third-party app development

SyntheticMass will provide a sandbox for Health IT developers, researchers, and clinicians interested in exploring new healthcare solutions. It enables this through:

- Realistic data for fictional patients
- Data that is free of protected health information (PHI) and personally identifiable information (PII) constraints
- Datasets updated over time based on clinical healthcare models and epidemiological models of population health.

To contribute, please visit SyntheticMass on [Github](#)



About Synthea

Synthea – Synthetic Patient Population Simulator

[Synthea](#) is a Synthetic Patient Population Simulator that is used to generate the synthetic patients within SyntheticMass. Synthea outputs synthetic, realistic (but not real) patient data and associated health records in a variety of formats. Read our [wiki](#) for more information.

Currently, Synthea features:

- Birth to Death Lifecycle
- Configuration-based statistics and demographics (defaults with Massachusetts Census data)
- Modular Rule System
 - Drop in Generic Modules
 - Custom Ruby rules modules for additional capabilities
- Primary Care Encounters, Emergency Room Encounters, and Symptom-Driven Encounters
- Conditions, Allergies, Medications, Vaccinations, Observations/Vitals, Labs, Procedures, CarePlans
- Formats
 - FHIR (STU3 v1.8)
 - C-CDA
- Rendering Rules and Disease Modules with Graphviz

To contribute, please visit Synthea on [Github](#)



FHIR Highlights

Intuitive, easy to learn

EHR standard accepted **globally**

Truly **interoperable** – can be used on any system

Machine & **human** readable

Can **query** for specific data points

80 / 20 rule followed

Uses **RESTful APIs** to exchange data

Easier to **find & access** data

Integrates **diverse** data

Hierarchical metadata structure – **better organized**

Data organized as **Resources**

Resources can be **combined**



BUT! Do not forget about **CCD** ... it is still most commonly used

Clinical Resources

General

AllergyIntolerance 3
Condition (Problem) 3
Procedure 3
ClinicalImpression 0
FamilyMemberHistory 2
RiskAssessment 1
DetectedIssue 1

Care Provision

CarePlan 2
Goal 2
ReferralRequest 1
ProcedureRequest 3
NutritionalOrder 3
VisionPrescription 1

Medication & Immunization

Medication 3
MedicationRequest 3
MedicationDispense 2
MedicationStatement 3
Immunization 3
ImmunizationRecommendation 1

Diagnostics

Observation 5
DiagnosticRequest 3
ProcedureRequest 3
Specimen 2
BodySite 2
ImagingStudy 3
ImagingManifest 1

<https://www.hl7.org/fhir/resourcelist.html>

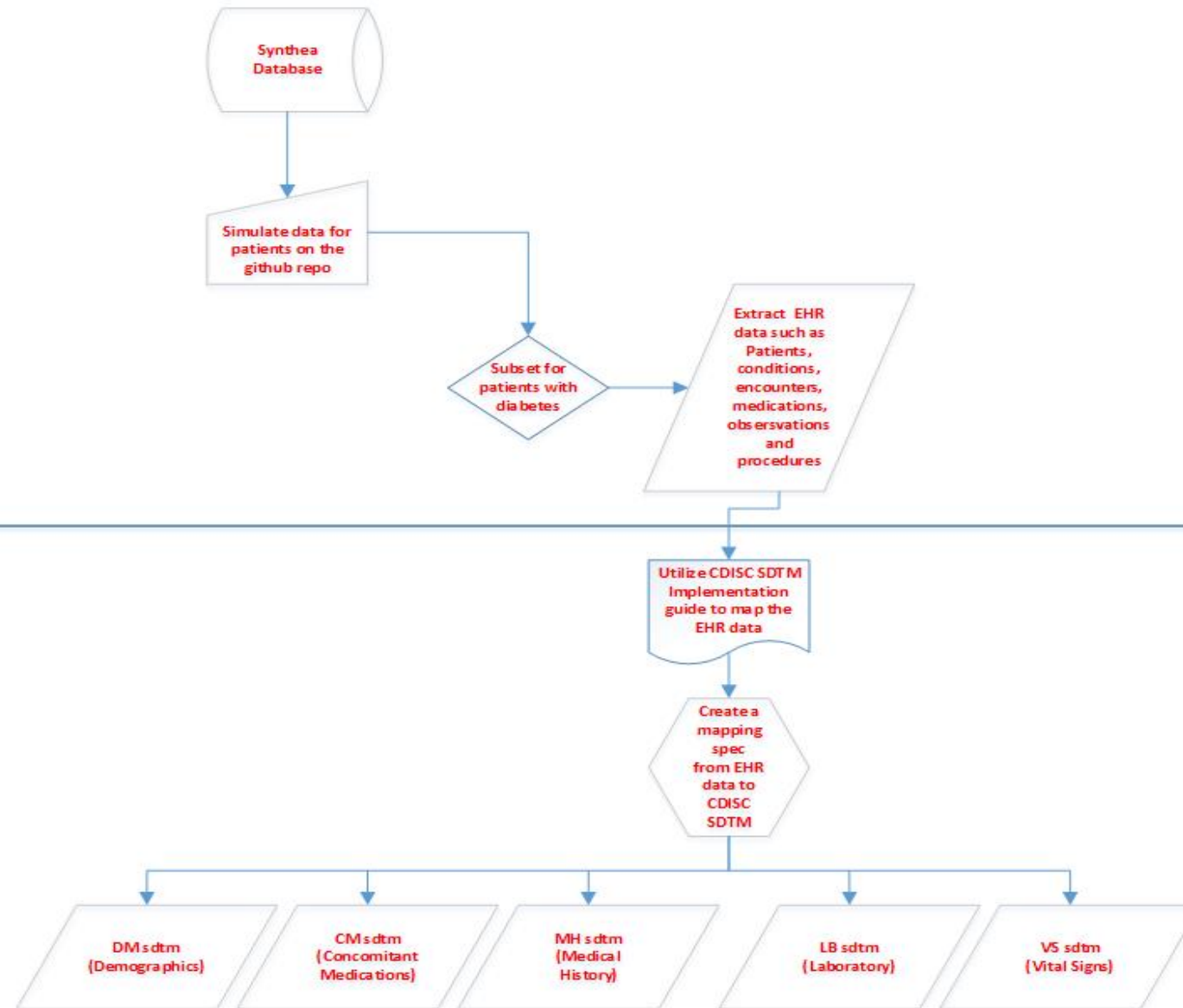
Method

Selection Criteria
Age >18 years of Age
Diagnosis of Type 2 Diabetes Mellitus (ICD-10 Code E11.n)
Subject has at least two encounters in the Synthea database
CDASH Data Fields Extracted
Unique identifier
Characteristics: Age, Sex, Weight, BMI
Concomitant Medications
Encounter Type
Labs: Hemoglobin A1C, Fasting Plasma Glucose, Creatinine (Diabetes TAUG)

EHR FHIR standards to CDISC SDTM standards process map

EHR FHIR standards

CDISC SDTM standards



Raw Data (JSON)

```
{
  "type": "collection",
  "entry": [
    {
      "fullUrl": "urn:uuid:54f8fc2f-5393-4cf7-939e-ec96c1a45732",
      "resource": {
        "id": "54f8fc2f-5393-4cf7-939e-ec96c1a45732",
        "meta": {
          "profile": [
            "http://standardhealthrecord.org/fhir/StructureDefinition/shr-demographics-PersonOfRecord"
          ]
        },
        "text": {
          "status": "generated",
          "div": "<div>Generated by <a href=\\"https://github.com/synthetichealth/synthea\\">Synthea</a>. Version identifier: decb8ce28c9fec6d0b58bfd5d498c98adcebdbb</div>"
        },
        "extension": [
          {
            "url": "http://hl7.org/fhir/StructureDefinition/us-core-race",
            "valueCodeableConcept": {
              "coding": [
                {
                  "system": "http://hl7.org/fhir/v3/Race",
                  "code": "2106-3",
                  "display": "White"
                }
              ]
            },
            "text": "race"
          }
        ]
      }
    }
  ]
}
```

Raw Data

ID	BIRTHDATE	DEATHDATE	SSN	DRIVERS	PASSPORT
a123c4e0-8618-416b-a489-440578e59bf1	3/13/1929		000 14 8006	000082252	Y21221458Y
bca189eb-7a41-4e42-86f4-4ee4d2aa5c1f	11/9/1986		999 56 4930	999975377	FALSE

PREFIX	FIRST	LAST	SUFFIX	MAIDEN	MARITAL	RACE	ETHNICITY	GENDER
Mrs.	Ellnore156	Toy711		Herman646	M	black	<u>dominican</u>	F
Mr.	Hilario17	Daugherty913			M	white	<u>german</u>	M

BIRTHPLACE	ADDRESS
Winchester MA US	278 Ibrahim Pike Dennis MA 02638 US
Westfield MA US	1486 <u>DuBuque</u> Turnpike Apt. 942 Fitchburg MA 01420 US

Mapping FHIR to CDASH and SDTM

Domain	Concept	FHIR RESOURCE PATH (<u>Resource.Element.</u> <u>Subelement</u>)	CDASH Variable	SDTM Variable	Notes
DM	Unique Subject ID	Patient.Identifier.value <u>Patient.Identifier.system</u>	USUBJID	USUBJID	In FHIR a unique patient is identified by a patient ID and a system identifier
	Sex	Patient.Gender	SEX	SEX	When converting to CDISC, gender must be re-coded
	Age	n/a	AGE	AGE	In FHIR, age is calculated for birth date and a reference date
	Birth Date	Patient.birthDATE	BRTHDAT	BRTHDTC	Date/times in FHIR are stored in ISO 8601. Date/times in SDTM are also stored in ISO 8601 format, but typically constructed from CDASH date/time components



Show Annotations

Demographics

Subject ID

[SDTM: DM.SUBJID]
[CDASH: DM.SUBJID]
[FHIR: Patient.Identifier]

What is the sex of the subject?

- ☐ M
- ☐ F
- ☐ UN
- ☐ U

[SDTM: DM.SEX]
[CDASH: DM.SEX]
[FHIR: Patient.Gender]

[SDTM: DM.SEX]
[CDASH: DM.SEX]
[FHIR: Patient.Gender]

What is the subject's age?

[SDTM: DM.AGE]
[CDASH: DM.AGE]
[FHIR: N/A]

What is the subject's date of birth?

[SDTM: DM.BRTHDTC]
[CDASH: DM.BRTHDAT]
[FHIR: Patient.birthDate]

What is the race of the subject?

- ☐ ASIAN
- ☐ WHITE
- ☐ AMERICAN INDIAN OR ALASKA NATIVE
- ☐ BLACK OR AFRICAN AMERICAN
- ☐ NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER

[SDTM: DM.RACE]
[CDASH: DM.RACE]
[FHIR: Patient.extension.StructureDefinition.us-core-race]

[SDTM: DM.RACE]
[CDASH: DM.RACE]
[FHIR: Patient.extension.StructureDefinition.us-core-race]

SAS Dataset

Study Identifier	Domain Abbreviation	Unique Subject Identifier	Date/Time of Birth	Age	Age Units	Sex	Race	Ethnicity	Country
FHIR001	DM	a123c4e0-8618-416b-a489-440578e59bfl	1929-03-13	88	YEARS	F	BLACK OR AFRICAN AMERICAN	HISPANIC OR LATINO	USA
FHIR001	DM	bca189eb-7a41-4e42-86f4-4ee4d2aa5clf	1986-11-09	30	YEARS	M	WHITE	NOT HISPANIC OR LATINO	USA

Interesting Things To Note

- Many of the FHIR resources can be used in CDASH & SDTM, **but not all**
- Some CDASH & SDTM data will still need to be entered/derived
- EHRs typically use SNOMED, ICD-10 & LOINC
- It is possible to extract multiple SDTM variables from one LOINC code in LB & VS (--CAT, --TEST, --UNIT, --SPEC, --METHOD)
- CDASH/SDTM do not have specific variables for Encounter type – this addition should be considered
- **All concepts** have different variable names in FHIR and CDASH/SDTM
- Controlled terminology is typically not harmonized

The Future

- FHIR allows for data subscriptions, or data queries – these can be used to extract ONLY the data needed and for real-time data access
- The EHR-to-CDASH (E2C) team must continue mapping the remaining FHIR domains
- Real-time translation code is needed
- Additional CDASH/SDTM variables may be useful to store the original SNOMED/ICD-10 codes
- CDASH should consider FHIR resources in future versions
- Additional FHIR extensions to better support research
- This PhUSE Research on FHIR team has more work to do ... **so come join us**

Thank you

Please contact Trisha Simpson with any questions:
Trisha.Simpson@UCB.com