

Using Real World Data & Synthetic Data to Predict Study Outcomes, Submission Performance, and Market Effectiveness

José C. Lacal – TrialTwin
Stuart Malcolm – Veramed, Ltd

Summary

10-year Vision

- Revolutionize trial data management
- Combining Historical and Synthetic Data
- Accelerate Design, Analysis, Submission
- Full life cycle: Compound to Consumer
- Expanding internal access to data, tools
- Will release most pieces as Open Source
- Looking for Collaborators

PHUSE Open Data Repository | 30+ Years Historical Data

Literature

Clinical
Trials

Regulatory
Approval

Adverse
Events

Vendor
Payments

Medicare
Payments

Federal
Payments

U.S. Department of Health & Human Services

NIH U.S. National Library of Medicine

NIH U.S. National Library of Medicine
ClinicalTrials.gov

EU Clinical Trials Register



Drugs@FDA: FDA Approved Drug Products



510(k) Clearances

FDA Adverse Event Reporting System (FAERS):
Latest Quarterly Data Files

MAUDE - Manufacturer and User Facility Device
Experience

How Open Payments Works



[Medicare Provider Utilization and Payment Data: Physician and Other Supplier](#)

[Medicare Provider Utilization and Payment Data: Inpatient](#)

[Medicare Provider Utilization and Payment Data: Outpatient](#)

[Medicare Provider Utilization and Payment Data: Part D Prescriber](#)

[Medicare Provider Utilization and Payment Data: Referring Durable Medical Equipment, Prosthetics, Orthotics and Supplies](#)

[Medicare Provider Utilization and Payment Data: Home Health Agencies](#)

USASPENDING.gov

FY 2018 OBLIGATED AMOUNT
\$6.6 Trillion
Data as of September 30, 2018

Chemical
Compound

PubChem

Using Open Data allows users to trace a drug's entire lifecycle:

- * starting with chemical compounds (NLM's PubChem)
- * through clinical trials (ClinicalTrials.gov, WHO's ITPR)
- * documentation on regulatory pathway (IND, NDA, etc.)
- * reported adverse events (FDA's FAERS)
- * manufacturer payments to providers (HHS' OpenPayments)
- * Medicare reimbursement data (CMS' Provider Utilization and Payment Data)

Synthetic Health Data | On-demand, At Scale, No Restrictions

Table: Synth_PHR_Demographics														
	id	Patient_ID *	First_Name	Middle_Name	Last_Name	Gender	Race	Date_Of_Birth	Country_Of_Birth	State_Province_Of_Birth	Location_Of_Birth	Location_Of_Birth_Lat	Location_Of_B	
	Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter	
1	603	0017d013-...	NAUSHIR	KONGSAB	KARIKEHALLI	M	ASIAN	2003-02-09	United States	Utah	College Ward	41.675876	-111.8779443	
2	768	003fafdd-...	JAMARL	HAN-SOO	TENNIAL	M	WHITE	1971-04-18	United States	Illinois	Johnsonville	40.212592	-87.6682441	
3	130	00c00d07-...	BONGJOON	JINCE	CADEGAN	M	NOT REPORTED	1975-04-05	United States	West Virginia	Gardner Junction	37.3934516	-81.0866469	
4	941	00e22265-a412-4cd8...	CHENG CHENG	SIRAH THOMAS	KINAMI	M	NOT REPORTED	2002-05-17	United States	Kentucky	Fariston	37.071089	-84.0560974	
5	974	0156c0cf-f2e8-4c37-...	NA-HYEON	DEOCALLI	MANOUSELIS	F	AMERICAN INDIAN OR ...	1960-07-31	United States	Ohio	Morristown	39.4125158	-82.1199232	
6	904	01b27b11-...	HEIJIN	VICKYANN	ALBETE	F	AMERICAN INDIAN OR ...	1965-09-16	United States	South Carolina	College Heights	34.3782099	-80.0594003	
7	280	02b671e7-316c-4928...	IQUBAL	KERSHAW	OWUSU-MCKENZIE	M	UNKNOWN	1961-12-11	United States	Georgia	Lenox	31.2725768	-83.4623845	
8	805	02c2be86-f1a9-452a...	VALLEY	JOHN PHILAMER	CHALVADAKIS	M	NOT REPORTED	1985-05-29	United States	Indiana	Mount Auburn	39.8123242	-85.1891872	
9	741	02e07c7f-36f9-4429-...	ZERLY	TAYLIAH	LABABIDY	F	NOT REPORTED	1995-07-23	United States	Nebraska	Page	42.4003369	-98.4153022	
10	142	03062ee7-9c41-4ef0-...	ZEDEENA	MAJO	NARDUCCIE	F	NOT REPORTED	2001-02-25	United States	Montana	Ridgelawn	47.7919621	-104.0868229	
11	516	030b93ec-...	ANANIAS	JIVONY	ZESZUT	M	ASIAN	1967-11-11	United States	Ohio	Shobers Mills ...	0.001	0.002	
12	844	03873d17-7f8b-4386...	MAATI	SUDHARMA	ROBILOTTA	F	UNKNOWN	1947-02-09	United States	South Dakota	Huron Colony	44.5644242	-98.2062586	
13	594	039b2d7a-...	SELORM	BIBINAZ	NIMOITYN	F	BLACK OR AFRICAN ...	1997-02-12	United States	Florida	Pompano Park	26.1982502	-80.2083805	
14	968	03ec43ab-bdf4-4172...	WILLIAM (BILL)	SHAKELL	PASTORIZA	M	BLACK OR AFRICAN ...	1968-12-18	United States	Washington	Chillowist	48.303039	-119.6355008	
15	645	049bc4fa-...	XAYDEN	TYQUINE	DING	M	BLACK OR AFRICAN ...	1949-11-02	United States	Mississippi	Rolling Heights ...	30.4437752	-88.9137826	
16	445	04beecb0-794e-48fb...	MALEEA	AMAIRANY	TOOMIN	F	BLACK OR AFRICAN ...	1948-08-03	United States	Arizona	Cheyenne Village	33.5980403	-112.136208	
17	847	0501f2ab-...	ARWIN ANGELAE	TESEN	RICCIO-HUCK	F	AMERICAN INDIAN OR ...	1986-12-03	United States	Virginia	Ardmore	37.3461907	-79.9353175	
18	82	0504391e-...	XUYEN	RHUSHI	JACKSON GRUEB	M	NATIVE HAWAIIAN OR ...	1942-05-10	United States	Texas	North Uvalde ...	29.2371272	-99.7997795	
19	1000	05462f77-6bba-4616...	JITRUDEE	RENITRA	BENUTO	F	AMERICAN INDIAN OR ...	1939-02-26	United States	Florida	Mandolin	28.0644	-82.615	
20	595	05578fdc-...	ROOLS	ATAL	FARNADY	M	UNKNOWN	1968-07-30	United States	Texas	Mays Crossing	28.9458631	-98.0628945	
21	905	055d3abf-0767-413c-...	SRISUDA	SHERI-RAE	SCAVONE	F	BLACK OR AFRICAN ...	1984-04-29	United States	Washington	Mead	47.7648923	-117.353439	
22	923	05849699-...	DY	HOUKUN	LADOVSKY	M	BLACK OR AFRICAN ...	1961-09-04	United States	Delaware	McDonalds ...	38.6651139	-75.4489203	
23	626	05d75b47-1578-4b10...	JERELL	VELDIN	GLANTON-...	M	WHITE	1952-08-27	United States	Michigan	Riverview Estat...	44.891	-83.4347222	
24	807	05dc78fb-7d95-4fba-...	MOMOLU	BYUNG SOO	MABEN-FAESTER	M	ASIAN	1943-07-31	United States	Alabama	Philadelphia	31.1615134	-85.1054299	
25	78	05de6106-39ef-40bd...	KYUNGMI	RATONA	LOEBL	F	NOT REPORTED	1975-03-10	United States	Florida	Azalea Park	28.5406128	-81.2991237	
26	70	05fe9ffa-...	YATISH	MOHAMMED ...	AGBI	M	WHITE	1962-10-07	United States	Texas	Carr	28.7343267	-96.7539808	
27	244	06551305-bf12-43cc...	CHRIS	YUNSUK	MAYMIND	M	ASIAN	1990-02-15	United States	Minnesota	Tyler	44.2772995	-96.132255	
28	406	0660e0ed-b697-4bcd...	DAINELL	ANYLLAH	KUBOJIRI	F	UNKNOWN	1938-11-02	United States	Kentucky	Jones (historical)	37.3619805	-83.2323765	
29	386	0679cdf-fef6a-45bd-...	DAYNA CARMEN	WASHAUNDR	BUSICO	F	BLACK OR AFRICAN ...	1941-11-02	United States	Tennessee	Blount Beach ...	35.7859368	-83.8851344	

Realistic yet
fake full
name

User-defined
age range, and
gender, race %.

Real city / town.
With Latitude,
Longitude
for mapping.

SD1629 | RWD +
Synth Data =>
Predictions |
USConnect23 4

Synthetic Health Data | On-demand, At Scale, No Restrictions

Table: Synth_PHR_Conditions										
	id	Patient_ID	Event_Date	TermID	TermName	UMLS_CUI	Occurrence	Usage	NIHPO_Hierarchy	
	Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter
1		1	f4ccc262-83ce-4580-8f47-b23f3ee1eecf	1979-09-12	73181007	Recurrent pterygium (disorder)	C0155158	1.0	0.0004	disorder
2		2	f4ccc262-83ce-4580-8f47-b23f3ee1eecf	2019-07-24	70637004	Cellulitis of toe (disorder)	C0263134	2.0	0.0032	disorder
3		3	f4ccc262-83ce-4580-8f47-b23f3ee1eecf	1994-09-09	59214008	Reduction mammoplasty (procedure)	C0191922	1.0	0.0003	procedure
4		4	f4ccc262-83ce-4580-8f47-b23f3ee1eecf	2017-12-25	370143000	Major depressive disorder (disorder)	C1269683	3.0	0.0228	disorder
5		5	f4ccc262-83ce-4580-8f47-b23f3ee1eecf	2010-10-14	304026004	Varicose ulcer of lower extremity (disorder)	C0553570	2.0	0.0021	disorder
6		6	47d3ac87-57a7-4f37-a1cb-1c6ce74fcfc	1961-07-29	389026000	Ascites (disorder)	C0003962	7.0	0.0283	disorder
7		7	47d3ac87-57a7-4f37-a1cb-1c6ce74fcfc	2020-04-07	274769005	Albuminuria (finding)	C0001925	2.0	0.003	finding
8		8	47d3ac87-57a7-4f37-a1cb-1c6ce74fcfc	1996-08-19	62479008	Acquired immune deficiency syndrome (disorder)	C0001175	6.0	0.0304	disorder
9		9	47d3ac87-57a7-4f37-a1cb-1c6ce74fcfc	1999-03-02	26947005	Open wound of lower limb (disorder)	C0178323	2.0	0.0027	disorder
10		10	47d3ac87-57a7-4f37-a1cb-1c6ce74fcfc	1948-04-26	1055001	Stenosis of precerebral artery (disorder)	C0265089	1.0	0.0022	disorder
11		11	21b07450-c530-49cb-a5f3-144031eedf90	1959-01-30	48167000	Amnesia (finding)	C0002622	6.0	0.0597	finding
12		12	21b07450-c530-49cb-a5f3-144031eedf90	1972-08-04	63079007	Closed traumatic dislocation of hip joint (disorder)	C0434662			disorder
13		13	21b07450-c530-49cb-a5f3-144031eedf90	2006-04-17	34068001	Heart valve replacement (procedure)	C0190173	4.0	0.002	procedure
14		14	21b07450-c530-49cb-a5f3-144031eedf90	1944-01-26	91857003	Acute lymphoid leukemia,disease (disorder)	C0023449	4.0	0.0059	disorder
15		15	21b07450-c530-49cb-a5f3-144031eedf90	2005-06-27	254588001	Polyp of intestine (disorder)	C0021846	1.0	0.0004	disorder
16		16	9460981d-e87d-4c26-aedd-e811d08eb00e	1973-08-20	235159007	Percutaneous endoscopic insertion of gastrostomy tube ...	C0176751	2.0	0.0009	procedure
17		17	9460981d-e87d-4c26-aedd-e811d08eb00e	1967-09-09	48194001	Pregnancy-induced hypertension (disorder)	C0852036	4.0	0.0225	disorder
18		18	9460981d-e87d-4c26-aedd-e811d08eb00e	1963-07-07	60826002	Coccidioidomycosis (disorder)	C0009186	1.0	0.0006	disorder
19		19	9460981d-e87d-4c26-aedd-e811d08eb00e	1988-11-16	36995006	Drainage of external ear (procedure)	C0198018	1.0	0.0003	procedure
20		20	9460981d-e87d-4c26-aedd-e811d08eb00e	1967-02-06	202482009	Wrist joint pain (finding)	C0221785	1.0	0.0009	finding
21		21	15df90b8-2d57-49d1-8ad5-600f695c64b5	1990-12-08	53627009	Closed fracture of radius AND ulna (disorder)	C0272624	1.0	0.0034	disorder
22		22	15df90b8-2d57-49d1-8ad5-600f695c64b5	2009-04-10	429699009	History of malignant neoplasm of colon (situation)	C1998265	5.0	0.0232	situation
23		23	15df90b8-2d57-49d1-8ad5-600f695c64b5	2011-10-26	266609001	Female infertility of anovulatory origin (disorder)	C0404572	2.0	0.0009	disorder
24		24	15df90b8-2d57-49d1-8ad5-600f695c64b5	2016-09-07	206118007	Fetal or neonatal effect of disproportion during labor and ...	C0473841	1.0	0.0061	disorder
25		25	15df90b8-2d57-49d1-8ad5-600f695c64b5	2021-01-16	236443009	Medullary sponge kidney (disorder)	C0022681	1.0	0.0005	disorder
26		26	8bc6fa31-57c2-4dd5-b600-16e29dd7728d	1997-11-21	6240004	Operative procedure on knee (procedure)	C0187769	1.0	0.0007	procedure
27		27	8bc6fa31-57c2-4dd5-b600-16e29dd7728d	2014-09-12	92258000	Benign neoplasm of orbit (disorder)	C0154023	1.0	0.0005	disorder
28		28	8bc6fa31-57c2-4dd5-b600-16e29dd7728d	1988-03-14	242489002	Accident due to contact with hot or corrosive substance ...	C0417508	1.0	0.0005	event

SNOMED-CT
codes

Full concept
description

UMLS CUI code
included as well

SD1629 | RWD +
Synth Data =>
Predictions |
USConnect23 5

Lifecycle of Medical Device or Pharma Product

Lifecycle

- Market Research: what disease / condition to target?
- Novel Compound: substance with expected positive impact
- Regulatory Submission: data submitted to seek clearance
- Approval: regulatory clearance to market in target jurisdiction
- Withdrawal / recall: end of the device's / drug's usable life



Predicting Outcomes | 10-Year Vision

Prediction Level 01

Simulate the full regulatory submission package from the chemical structure of the new compound.

- Historical data of chemical structures cleared earlier
- Synthetic data of Subjects. Even before FPI.

Prediction Level 02

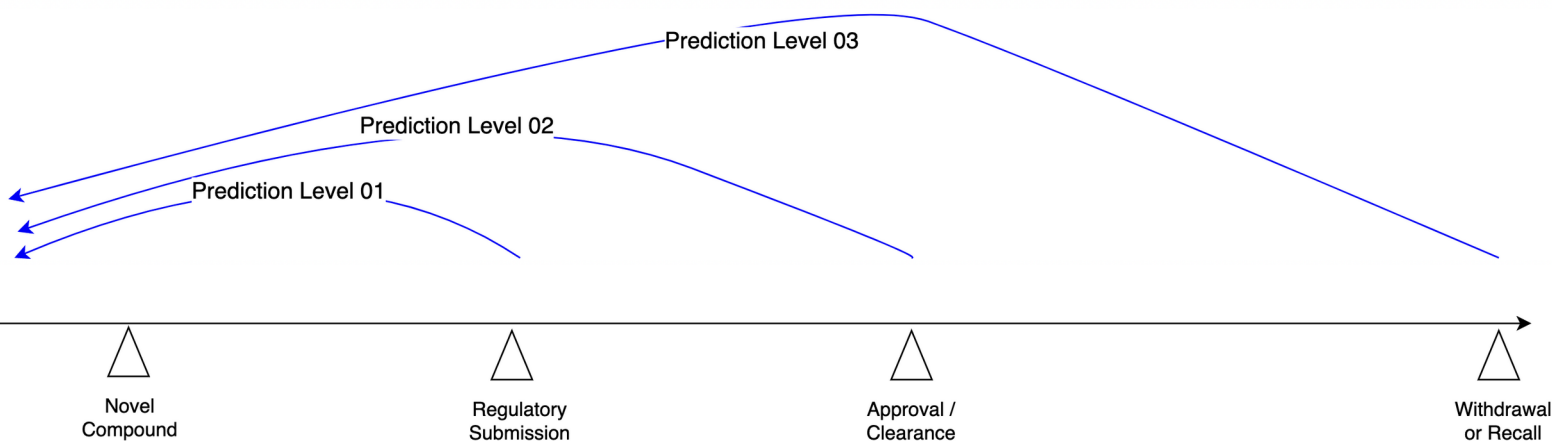
Simulate taking a new compound all the way out to regulatory approval.

- Historical data of approval pathway followed by previous devices / drugs in similar class.

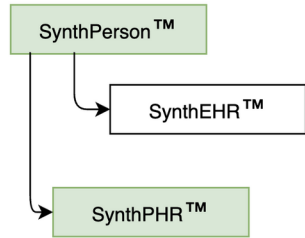
Prediction Level 03

Simulate the entire life-cycle of a new compound out to the marketing and revenue-producing stage.

- Historical data of usage, reimbursement
- Synthetic simulation of target population



Phase I | *Build a Synthetic Population*



Synthetic Population

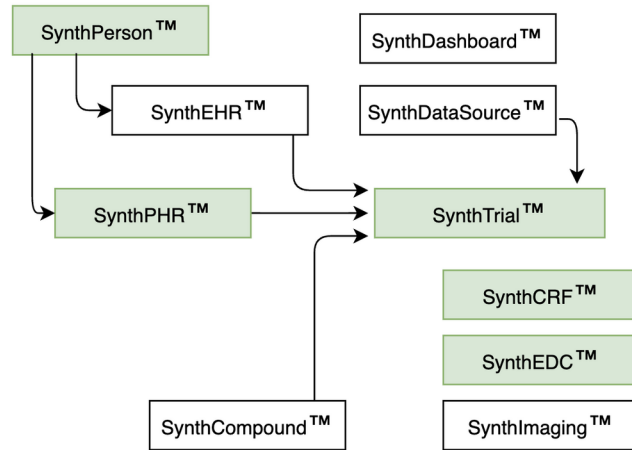
- Create scientifically-accurate Synthetic Persons.
- Create full Synthetic Personal Health Record for each Person:
 - conditions
 - devices
 - drugs
 - lab results
- Simulate hospital visits from Electronic Health Records



Phase II | *Run a Synthetic Clinical Trial*

Start Data Mgmt. Early

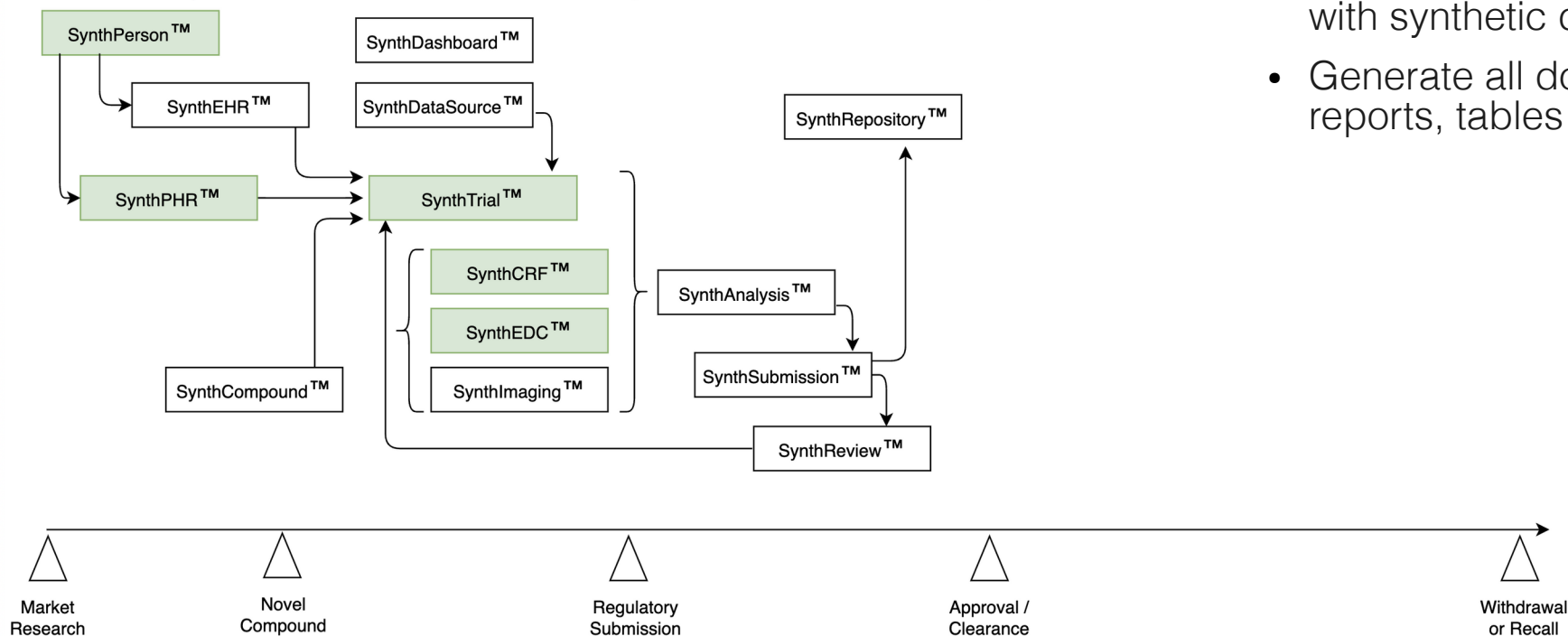
- Build study in EDC
- With trial-specific parameters
- Populate each CRF with realistic, fake synthetic data
- Simulate RWD coming from non-CRF data sources



Phase III | *Synthetic Submission*

Start Analysis Before FPI

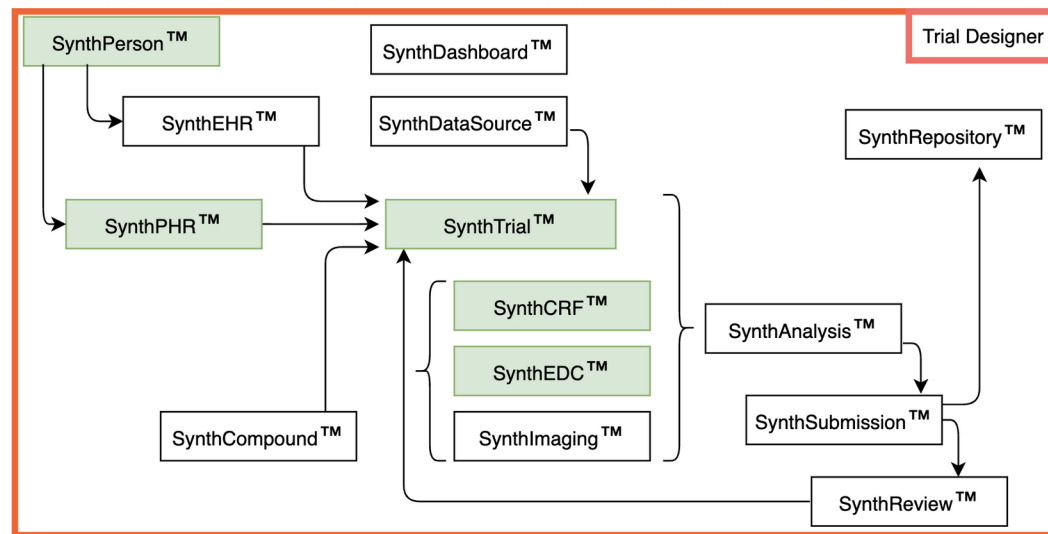
- Build, test, and validate entire analysis process early
- Prepare a test full submission with synthetic data
- Generate all documents, reports, tables and figures



Accelerating Analysis | *Trial Designer*

Integrated Simulation

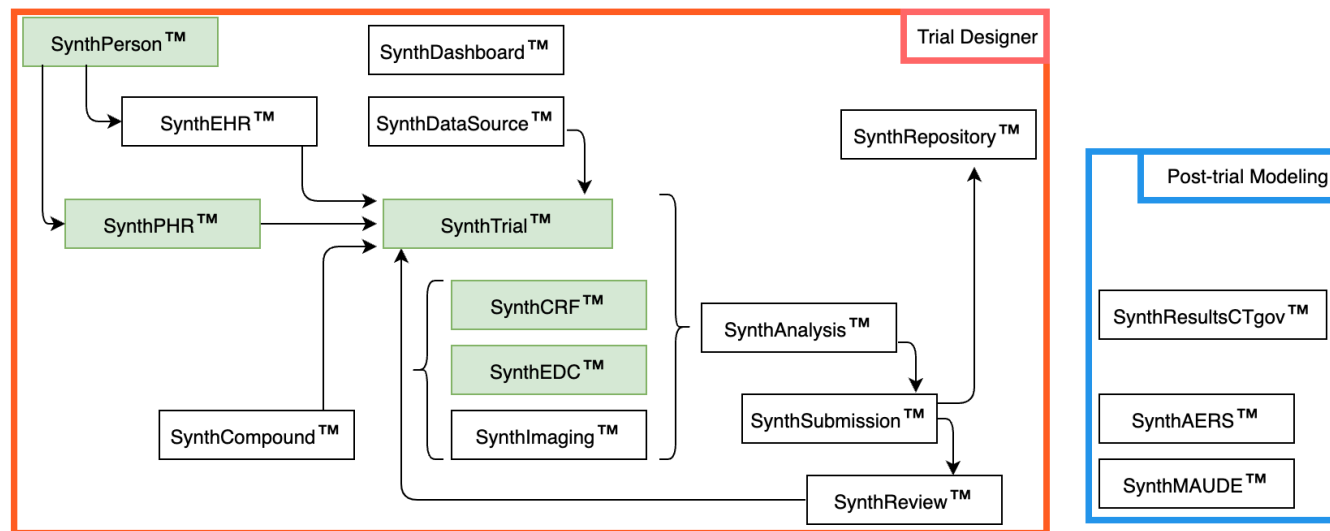
- Build, test, and analyze each trial from one User Interface
- Web-based tool avoids vendor lock-in, compatibility issues
- Different roles, user access
- Trial Designer to be available later this year (Open Source)



Phase IV | *Post-trial Modeling*

Predict Operating Life

- Use historical data to predict adverse events: device, drug
- Simulate uploading synth trial results to ClinicalTrials.gov



△
Market
Research

△
Novel
Compound

△
Regulatory
Submission

△
Approval /
Clearance

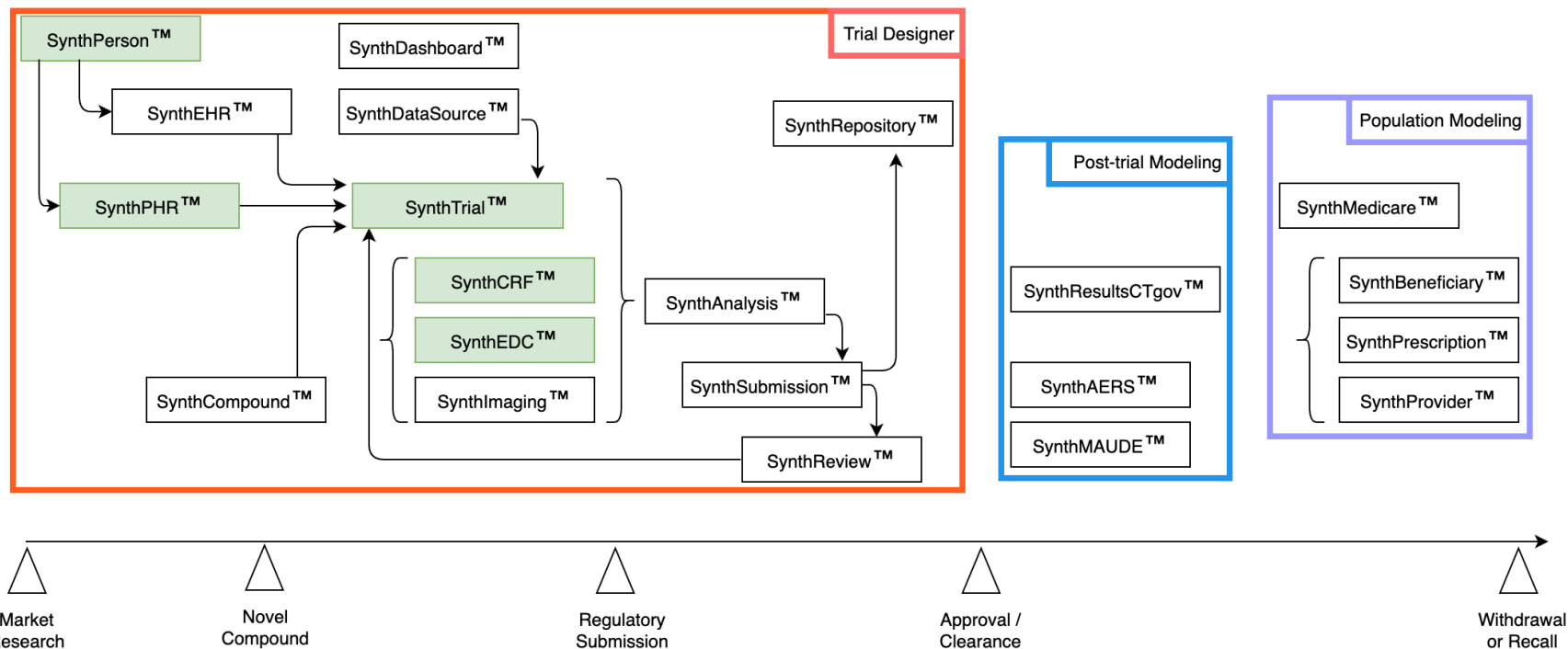
△
Withdrawal
or Recall

SD1629 | RWD +
Synth Data =>
Predictions |
USConnect23 12

Phase V | Population Modeling

Prescription Patterns

- Use historical data to model marketing pathways
- (US-specific) Medicare data



Looking for Collaborators

Fund or Provide Requirements

- Sponsors
- CROs
- Regulators

Contact Us

Co-authors

- José C. Lacal
Jose.Lacal@NIHPO.com
- Stuart Malcolm
Stuart.Malcolm@Veramed.co.uk