Developing Real-World Data ad Evidence to Support Regulatory Decision-Making

National Press Club

529 14th St NW, Washington, DC 20045

October 3, 2019

Welcome and Overview

Welcome and Update from FDA

Emerging Insights into the Development of RWE from Randomized Designs

Real World Evidence *with* Randomized Clinical Trials

Adrian F. Hernandez, MD, MHS
Vice Dean for Clinical Research
Duke School of Medicine

@texhern

What's a problem we're aiming to solve?

~2% **~90%**

McHugh, K., Swamy, G., & Hernandez, A. (2018). *Journal of Clinical and Translational Science, 2*(6), 384-392. doi:10.1017/cts.2019.1

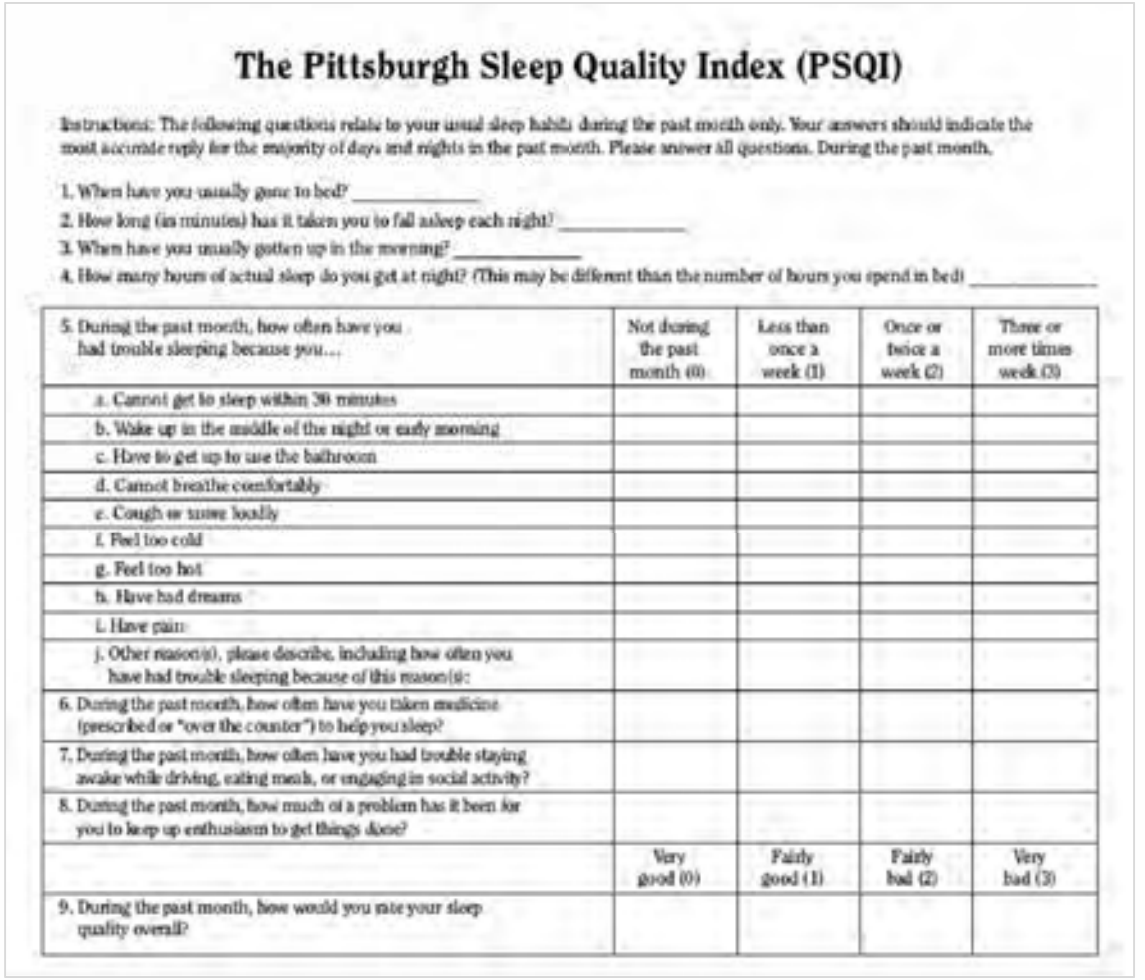
21,000

- Who are these pioneers?
- Why did they agree to participate?

Ideal Experience?

Why do people do it?

Traditional clinical studies feel like work.



The Pittsburgh Sleep Quality Index (PSQI)

Instructions: The following questions relate to your usual sleep habits during the past month only. Your answers should indicate the most accurate reply for the majority of days and nights in the past month. Please answer all questions. During the past month,

1. When have you usually gone to bed? _____
2. How long (in minutes) has it taken you to fall asleep each night? _____
3. When have you usually gotten up in the morning? _____
4. How many hours of actual sleep do you get at night? (This may be different than the number of hours you spend in bed) _____

5. During the past month, how often have you had trouble sleeping because you…	Not during the past month (0)	Less than once a week (1)	Once or twice a week (2)	Three or more times week (3)
a. Cannot get to sleep within 30 minutes				
b. Wake up in the middle of the night or early morning				
c. Have to get up to use the bathroom				
d. Cannot breathe comfortably				
e. Cough or snore loudly				
f. Feel too cold				
g. Feel too hot				
h. Have bad dreams				
i. Have pain				
j. Other reason(s), please describe, including how often you have had trouble sleeping because of this reason(s):				
6. During the past month, how often have you taken medicine (prescribed or "over the counter") to help you sleep?				
7. During the past month, how often have you had trouble staying awake while driving, eating meals, or engaging in social activity?				
8. During the past month, how much of a problem has it been for you to keep up enthusiasm to get things done?				
	Very good (0)	Fairly good (1)	Fairly bad (2)	Very bad (3)
9. During the past month, how would you rate your sleep quality overall?				

Yet, people want an experience like this...

Convenient
Flexible
Personalized

Hope for the real world?

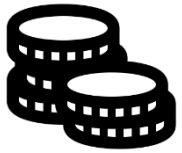
Health Systems Want Better Data

Diagnostics & Analytics

Safety and Harm Reduction

Preventive Health

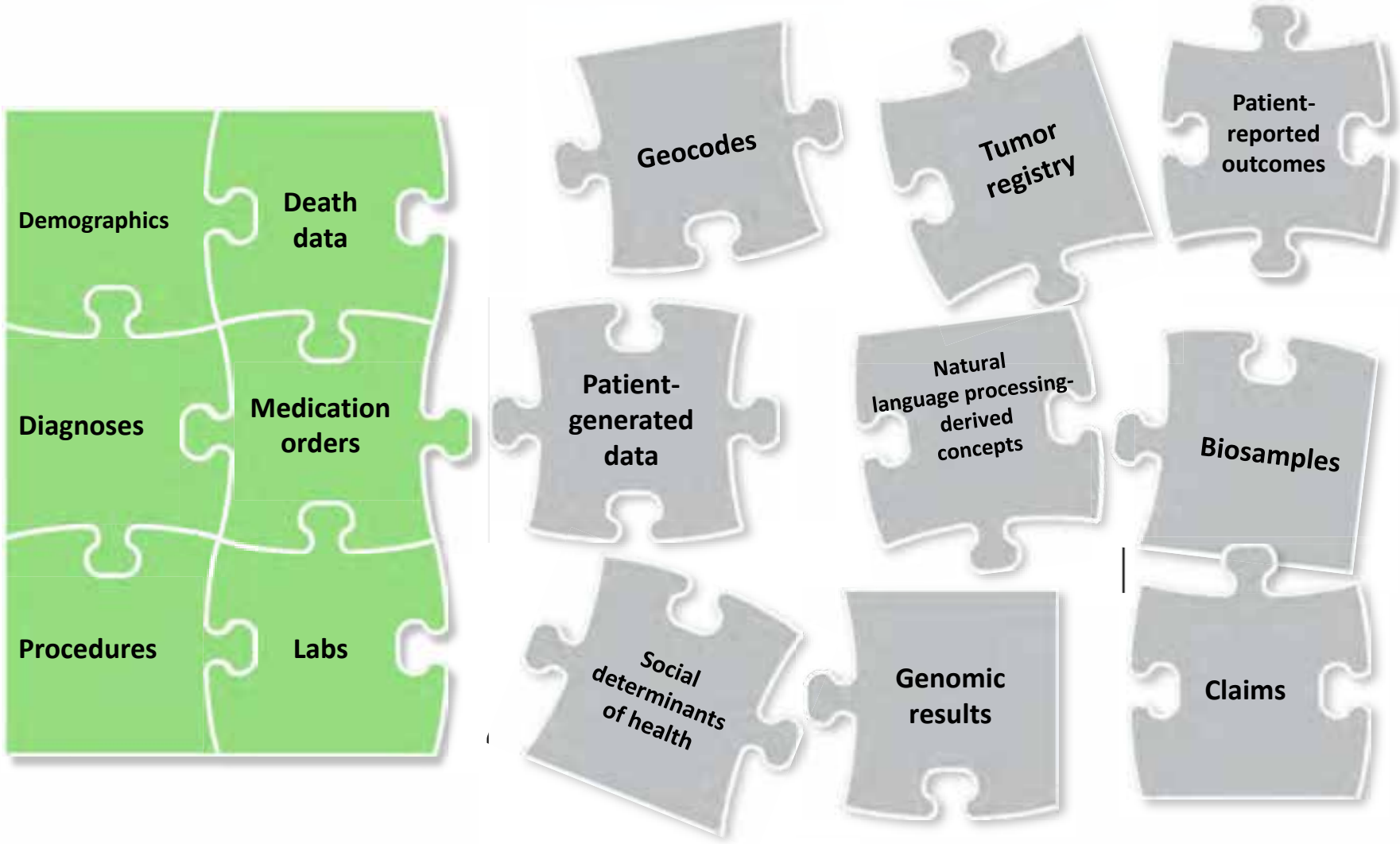
Precision Health



Cost Reduction

Population Health

#1 Data Everywhere & Curation



#2 People-Centeredness

Direct to Consumer

- Flexible
- Frictionless
- Fun

Direct to Participant

- Personalized
- Streamlined
- Valuable

PELOTON

#3 mHealth Technology

User-Reported Data

What people say

Task-Based Measures

Measures effort and physiology

Passive Sensing

What people actually do day to day

15

A Real World Example

Looking Back at a Disruptive Technology

EFFECTIVENESS OF INTRAVENOUS THROMBOLYTIC TREATMENT IN ACUTE MYOCARDIAL INFARCTION

GRUPPO ITALIANO PER LO STUDIO DELLA STREPTOCHINASI NELL'INFARTO MIOCARDICO (GISSI)*

Summary In an unblinded trial of intravenous streptokinase (SK) in early acute myocardial infarction, 11 806 patients in one hundred and seventy-six coronary care units were enrolled over 17 months. Patients admitted within 12 h after the onset of symptoms and with no contraindications to SK were randomised to receive SK in addition to usual treatment and complete data were obtained in 11 712. At 21 days overall hospital mortality was 10·7% in SK recipients versus 13% in controls, an 18% reduction (p=0·0002, relative risk 0·81). The extent of the beneficial effect appears to be a function of time from onset of pain to SK infusion (relative risks 0·74, 0·80, 0·87, and 1·19 for the 0–3, 3–6, 6–9, and 9–12 h subgroups). SK seems to be a safe drug for routine administration in acute myocardial infarction.

The Lancet · Saturday 22 February 1986

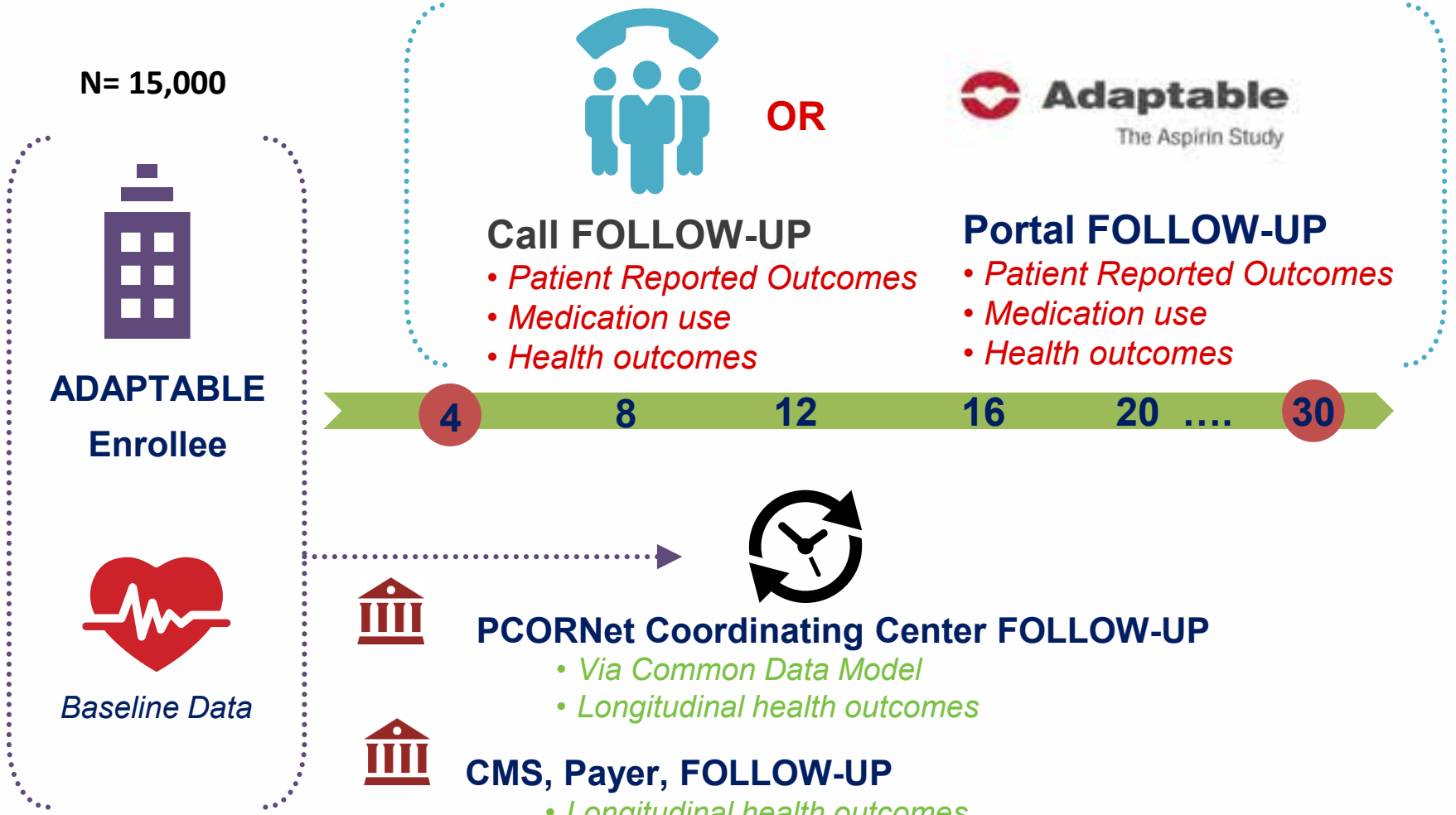
"It started with no funding and skepticism in some quarters but today GISSI is recognized as an Italian achievement that has changed cardiology treatment worldwide."

http://eurheartj.oxfordjournals.org/content/31/9/1023.full

17

ADAPTABLE: What's the Right Dose of Aspirin?

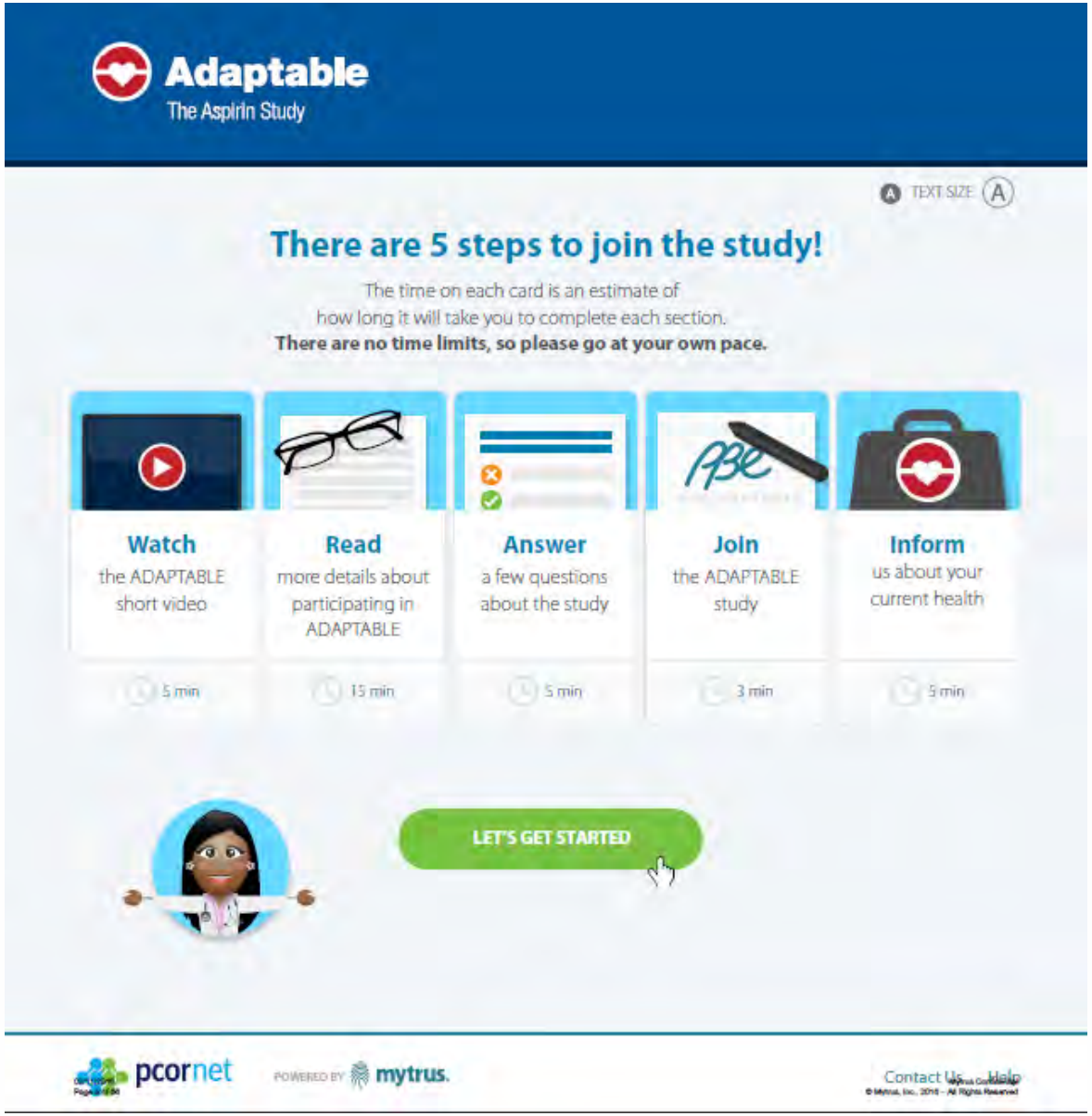
eScreening, eEnrollment and eFollow-up



http://adaptablepatient.com

18

The Participant Portal



adaptablepatient.com

Real World Evidence with Clinical Trials Roadmap

Engagement, Electronic Health Data and Embedded Delivery

Pre-study

- Assess sites' use of EHR to facilitate research
- Usability of inclusion and exclusion criteria
- Refine protocol
- Community interaction profiles with health system
- Feasibility analysis
- Recruitment plan

Study setup

- Utilize EHR to identify local participants
- Embed encounter instructions and site content into EHR
- Pre-consent & study specific consent
- Alert clinician about trial
- Model outcomes

Recruitment

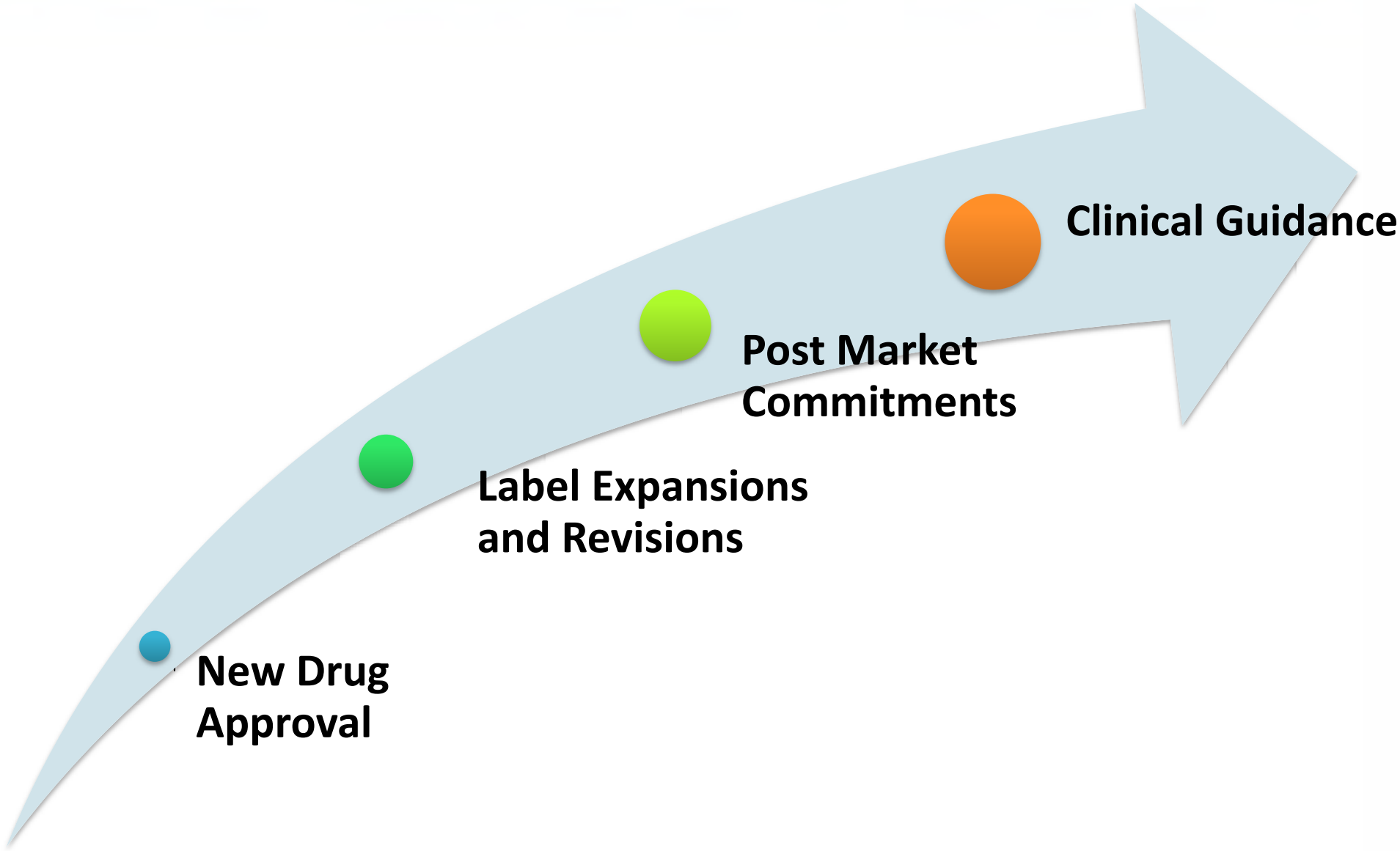
- E-consent with comprehension questions
- Incorporate screening criteria into EHR for
  - Scheduling patients
  - Contacting patients
  - Recruiting patients
- Alert clinician of patient eligibility
- EHR Health Portals
  - Patient opt in/out for types of studies

Study conduct

- Trials specific data capture from care delivery
- Auto-populated CRFs fields from EHR
- Extract data to facilitate work of study coordinator
- Query data to identify events
- Participant retention and education
- Return of results

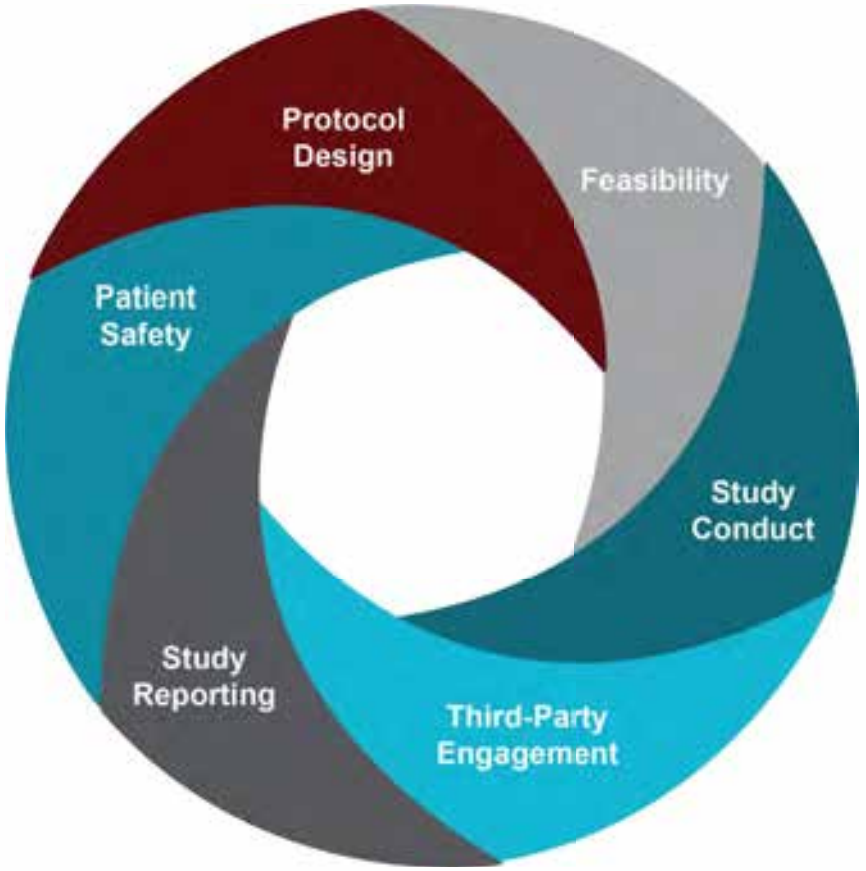
Important Matters
Quality & Outcomes

What's the Purpose?



Designing to the Purpose

Quality by Design



http://www.ctti-clinicaltrials.org/toolkit/QbD

Making Decisions:

Where do you fall the real world?

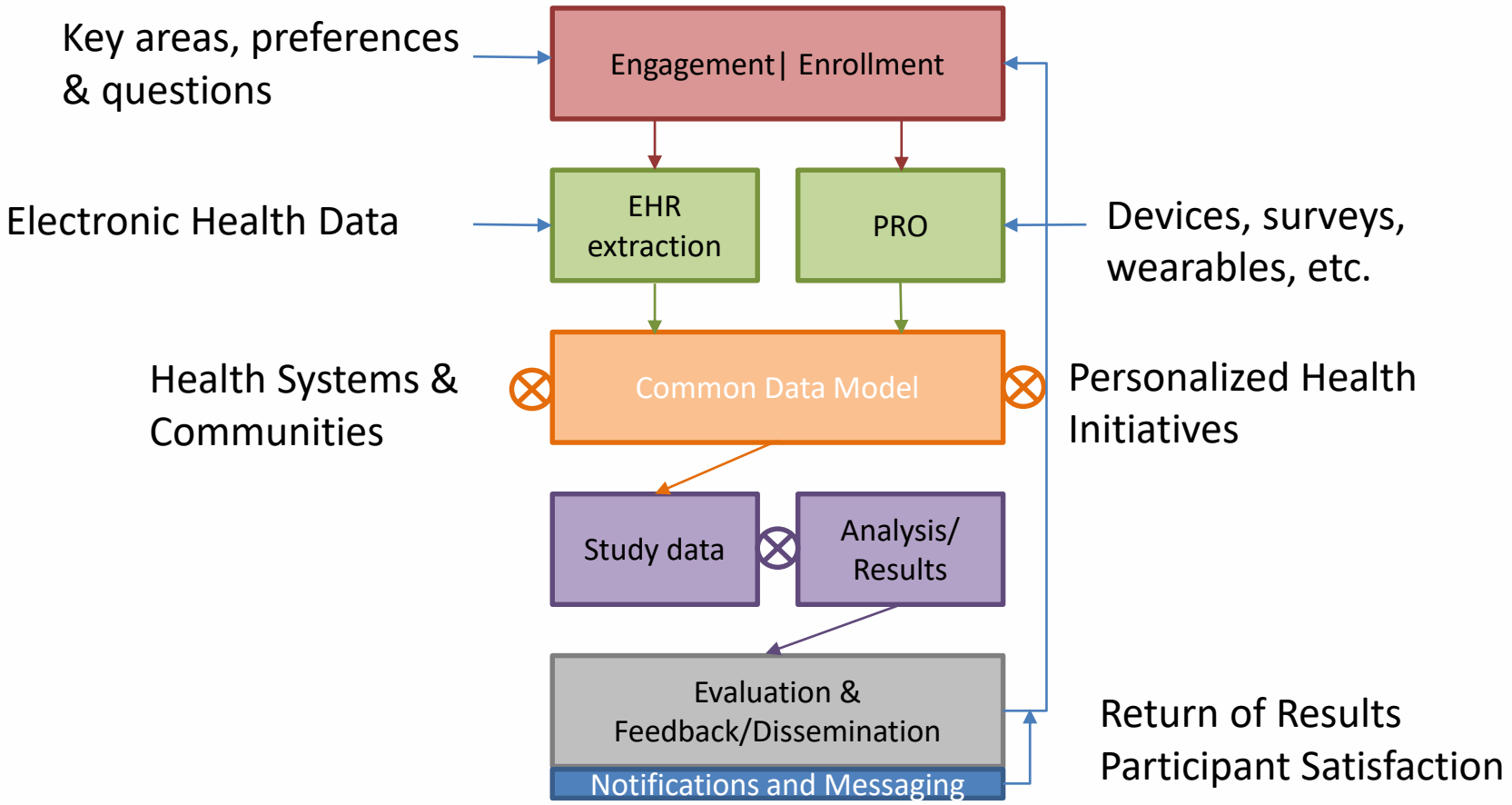
Ideal World

- Ideal Population
- Ideal/Perfect Care
- Blinding
- Placebo
- Coordinator Data Collection
- $$ is limitless

Real World

- Routine Population
- Usual Care
- [illegible]ed
- [illegible]ontrol
- Passive data collection
- Participant directed data collection
- $$ leveraged with embedded trials

The Puzzle Coming Together?



Participant Alumni Network
Clinician Engagement and Rejuvenation

Hernandez AF and Cruz H. Circ 2017

Emerging Real World Evidence

Match Unmet Needs with….

- Advances in curated health records
  - clinical, electronic health records, claims
- Advances in technology
- Advances in capturing digital exhaust
- Advances in phenotyping
- Advances in systems
- Advances in methods…**including randomized trials**

But to make this work we need…

“patient/clinician/system” engagement & trustworthy data

Emerging Insights into the Development of RWE from Randomized Designs

RCTs with Pragmatic Elements – Some Regulatory Considerations

October 3, 2019

Peter Stein, MD

Director

Office of New Drugs / CDER / FDA

A few comments on pragmatic randomized trials

Pragmatic trials: no standard definition - "explanatory" vs "pragmatic" *approaches* discussed by Schwartz and Lellouch (J Clin Epi 2009): biological assessment vs clinical relevance

- To support a **regulatory decision**, the issue is the **persuasiveness** of the findings to provide evidence of efficacy
- Randomization and blinding are **methods** to generate persuasive results
  - Randomization provides balance at treatment initiation
  - Blinding helps assure balance (of monitoring, adherence, endpoint assessment, continuation) *after* treatment initiation
- The "traditional" trial infrastructure is resource intensive and costly, *but*
  - Assures a patient population that is well defined, having the target condition
  - Provides careful, regular monitoring for collection of safety information and reliable trial endpoints
  - Has data that is well documented, stable, and traceable from source to results

How pragmatic a trial can be (and provide useful results), depends on the trial's purpose (e.g., regulatory, cost-effectiveness, comparative effectiveness, etc.) and the study question it seeks to answer

Regulatory "objectives": what key questions do we need clinical studies to answer?

- Does the drug *work* for the proposed indication?
  - Causal inference: *substantial evidence of effectiveness*
- Do the drug's benefits (clinical relevance of efficacy in the indicated patients) *outweigh* the drug's risks (expected or potential safety or tolerability concerns) in the indicated population (is it **safe** for use)?

Approvability

- Can we properly describe the dose/regimen, and the drug's safety profile and risks? (*Sections 2, 5, 6: D&A, W&P, Adverse Reactions*)
- Can we describe the supporting evidence from clinical trials (*Section 14: Clinical Studies*)?

Labeling

Pragmatic trials: two definitions and some questions

Pragmatic trials aim to determine **if an intervention works in real-world settings**, so that results can be **generalized to everyday practice** and support decision-making by patients, providers, and health system leaders; contrastingly, explanatory trials aim to determine **if and how an intervention works under well-defined and highly controlled conditions**

Taljaard et al. Trials 2018

Designed for the primary purpose of **informing decision-makers regarding the comparative balance of benefits, burdens and risks** of a biomedical or behavioral health intervention at the individual or population level.

Califf and Sugarman
Clinical Trials 2015

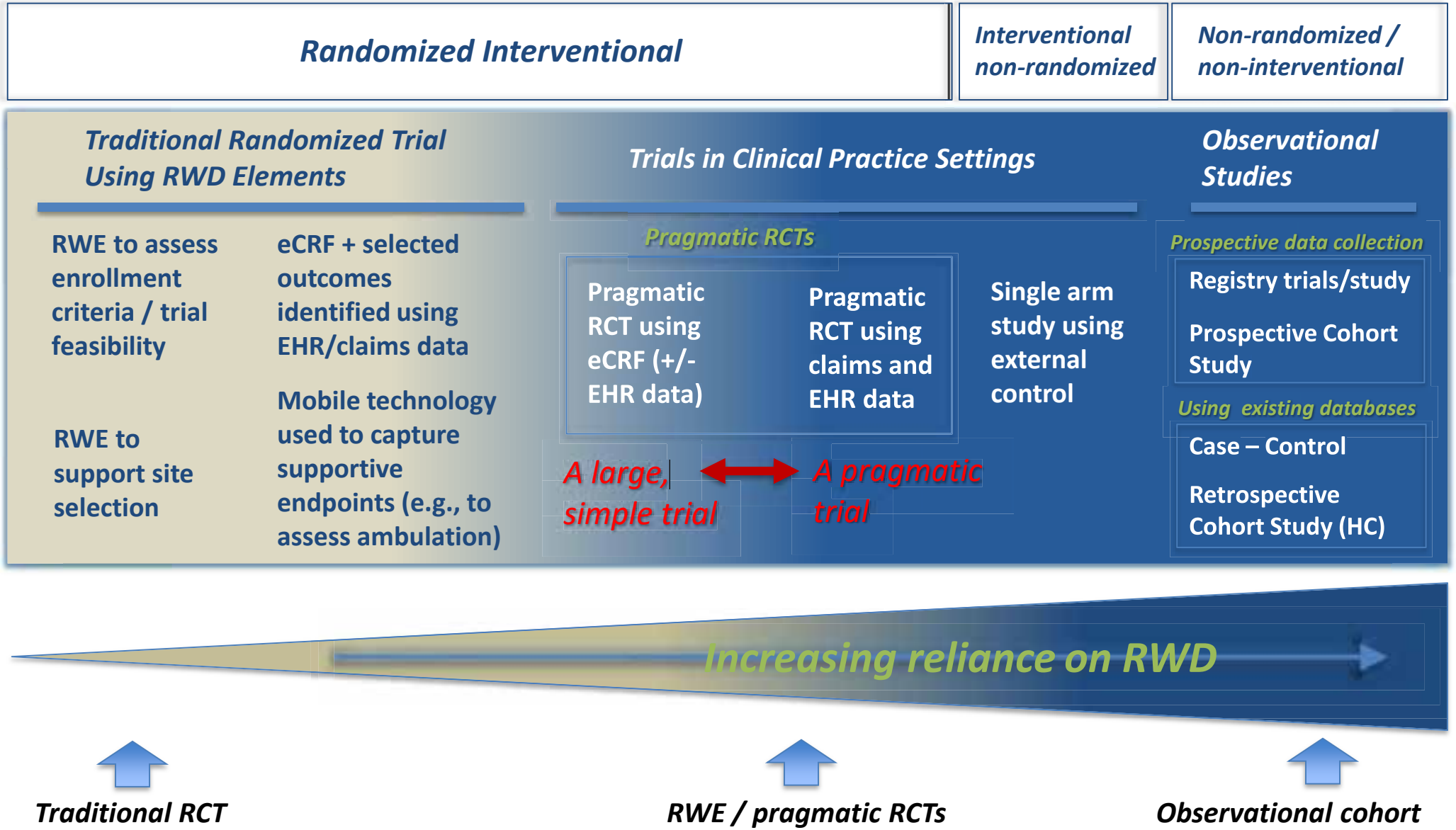
Issues raised:

- Assumes that "traditional" RCTs *do not* inform everyday practice – that results from such RCTs are *not generalizable*
  - What is the evidence for this?
  - What underlies differences in results between traditional RCTs and "pragmatic" trials?

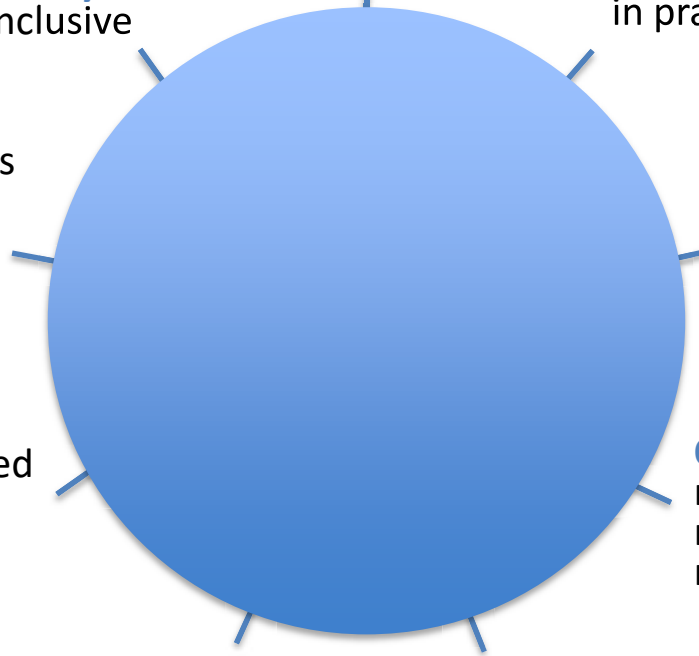
- When can trials with *less* well-defined and *less* well controlled conditions provide useful information?

- **Drug adherence**
- **Patient populations studied**
- **Interventions or co-interventions**
- **Monitoring**
- **Patient follow-up**
- **Endpoint assessment**
- **Data quality and reliability**

Wide spectrum of potential uses of RWD / RWE in clinical studies



Pragmatic randomized clinical trials: an overview of components



Study population: entry decision often by participating physician, few exclusions

Do the patients actually have the targeted disease?

Recruitment: patients in practice settings

Setting: typically community practices (general or specialty, but not usually at referral centers)

How well-controlled and reliable will patient monitoring and evaluation be?

Organization: often at sites not previously involved in research, usually limited or no research infrastructure

Intervention: usually not blinded; co-interventions not usually standardized/controlled

How important is blinding in supporting robust casual inferences?

Adherence: no specific efforts to assure higher adherence or to assess adherence (other than through claims for refills)

How much do we care to understand the effect if patients do not take the drug? Is adherence per se an issue?

Monitoring: may be no or limited protocol-defined requirements: follow-up as deemed clinically appropriate

How well are we detecting efficacy endpoints and safety?

Primary outcome: through claims or EHR, may use limited eCRF collection; often no required procedures; adjudication can be implemented (all or some)

How accurate and reliable is the endpoint, does it reflect what it purports to reflect?

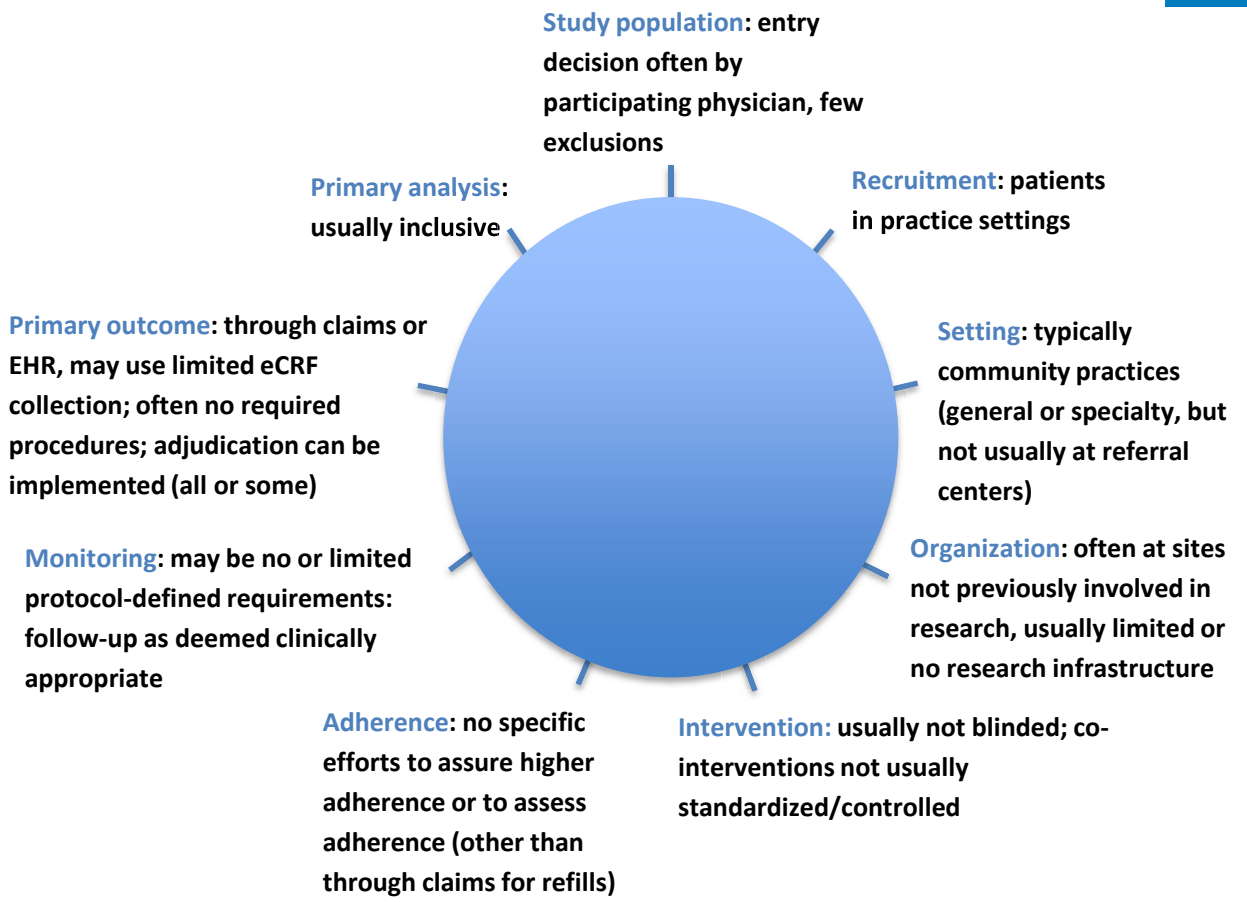
Primary analysis: usually inclusive

How well are the analysis populations constructed – do we understand the impact of missing data?

Based on: The PRagmatic-Explanatory Continuum Indicator Summary 2 (PRECIS-2) wheel BMJ 2015

Increasing use of trials with pragmatic feature(s)

- Identification of relevant questions for practitioners and patients
- Selection of an intervention *that can be appropriately delivered* in a clinical practice setting
- For studies of approved drugs, streamlined safety data collection
- Integration of clinical data across health care systems to maximize data capture
- If needed, utilize mobile technologies to fill in the gaps, including the capture of patient reported outcomes



Many trials can have 'pragmatic elements' while maintaining rigorous standards for data collection and assessment

Challenges of pragmatic trials

- **Design consistent with purpose** – if supporting regulatory decision-making, pragmatic elements may need to be balanced with elements assuring strong "believability"
- **Broader patient population** – but retaining minimum patient enrollment criteria to assure that the indicated population is studied
- **Interventions consistent with clinical practice** – but assuring patients get treatment to be studied (and adherence is evaluated)
- ***May* be unblinded** – but then need to have objective endpoints, consistent monitoring and balanced co-interventions
- **Meaningful endpoints -** that accurately evaluate study objective – whether using an eCRF or using EHR or claims data
- **Data that is reliable** - data (at least some) available for review, to assure accuracy of data, and fidelity of translation from source to analytic datasets
- **Patient follow-up sufficient** - assure that missingness (imbalanced, or informative) isn't confounding results

Emerging Insights into the Development of RWE from Randomized Designs

Session I: Establishing a High-Quality RWD Ecosystem

UCSF – Quantum Leap healthcare Collaborative

Integrating Clinical Care and Research

Adam Asare, PhD (UCSF, QLHC)
Laura Esserman, MD, MBA (UCSF, QLHC)
Mitra Rocca, PhD(FDA)
Sue Dubman, PhD (QLHC)

VISION: Integrate care process and research

OneSource

Data Entry

Patient-reported data:
- *From home or*
- *In clinic*

Clinician data entry
- *Structure forms*
- *EHR text notes*

Research Coordinator entry
- *CRFs*

Continuum of breast patient care

Data Uses

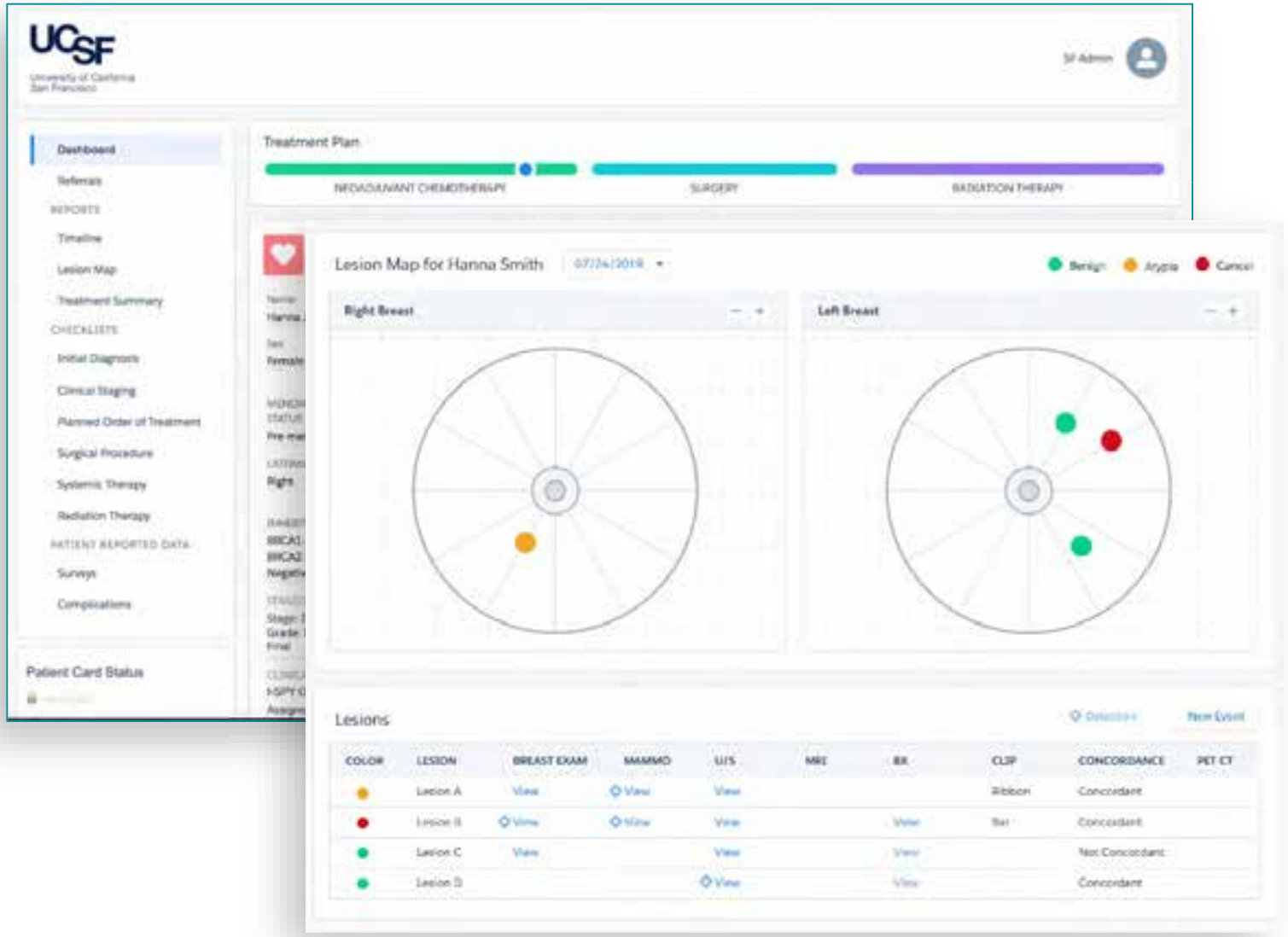
Dashboard / Reports
- *Summary dashboard*
- *Clinician Report*
- *Patient Report*
- *Tech Report*
- *Elevated Risk Report*
- *Tumor Board Report*
- *etc.*

Services / Referrals
- *Trial Matching*
- *Genetic Counseling*
- *Social Work*
- *Nutritionist*
- *Peer Support*
- *Behavioral / Sleep*
- *Psycho-Onc*
- *etc.*

Quality Improvement
- *Other trials, studies*
- *Registries*
- *etc.*

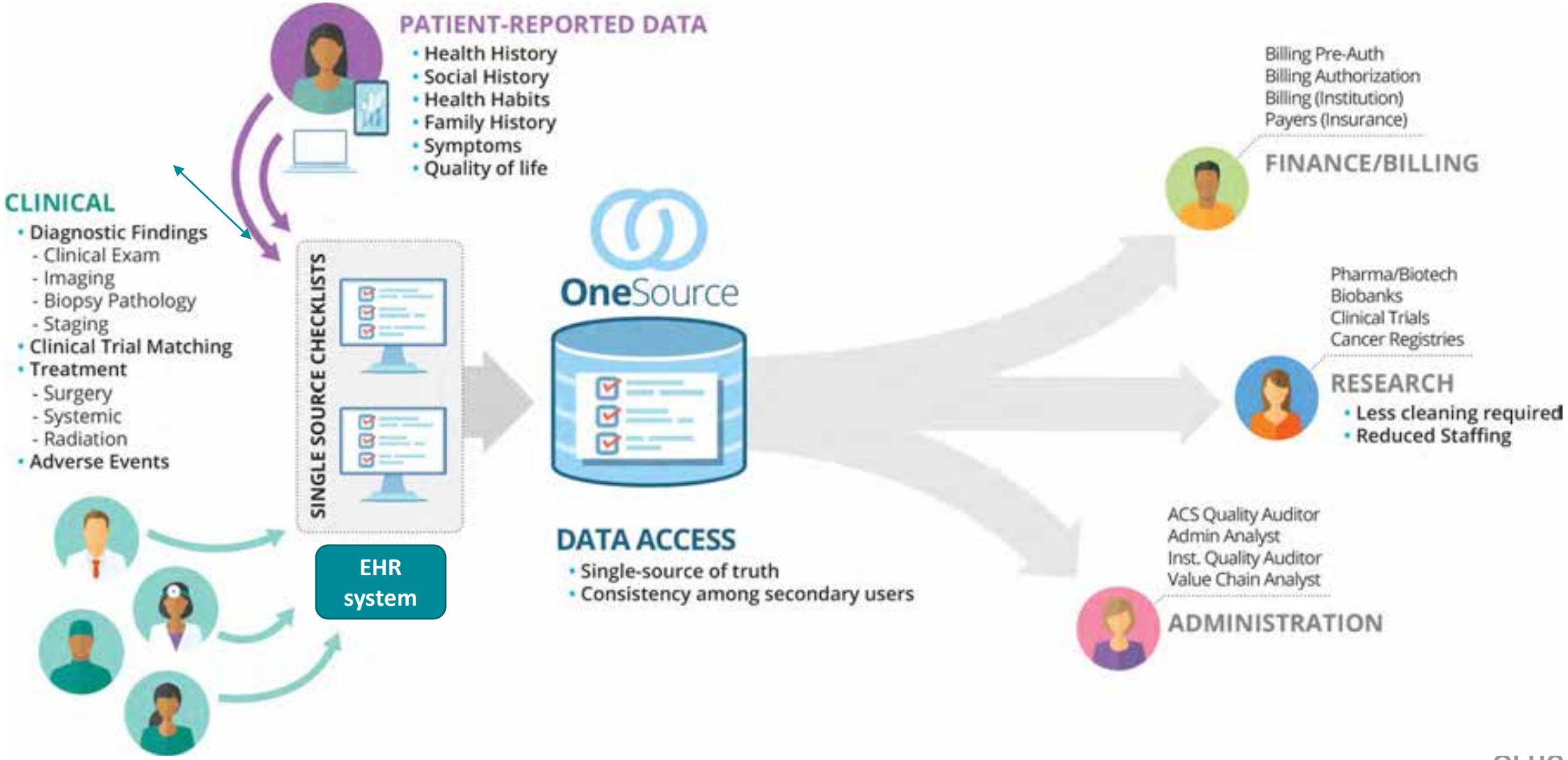
Decision support at point of care

- Centralizes and organizes trusted, structured data for clinical care,
- Provides for tailored decision support tools not readily supported by EHR systems
- Patients experience streamlined care delivered by empowered teams that are continuously learning and improving.



Structured data as “source”

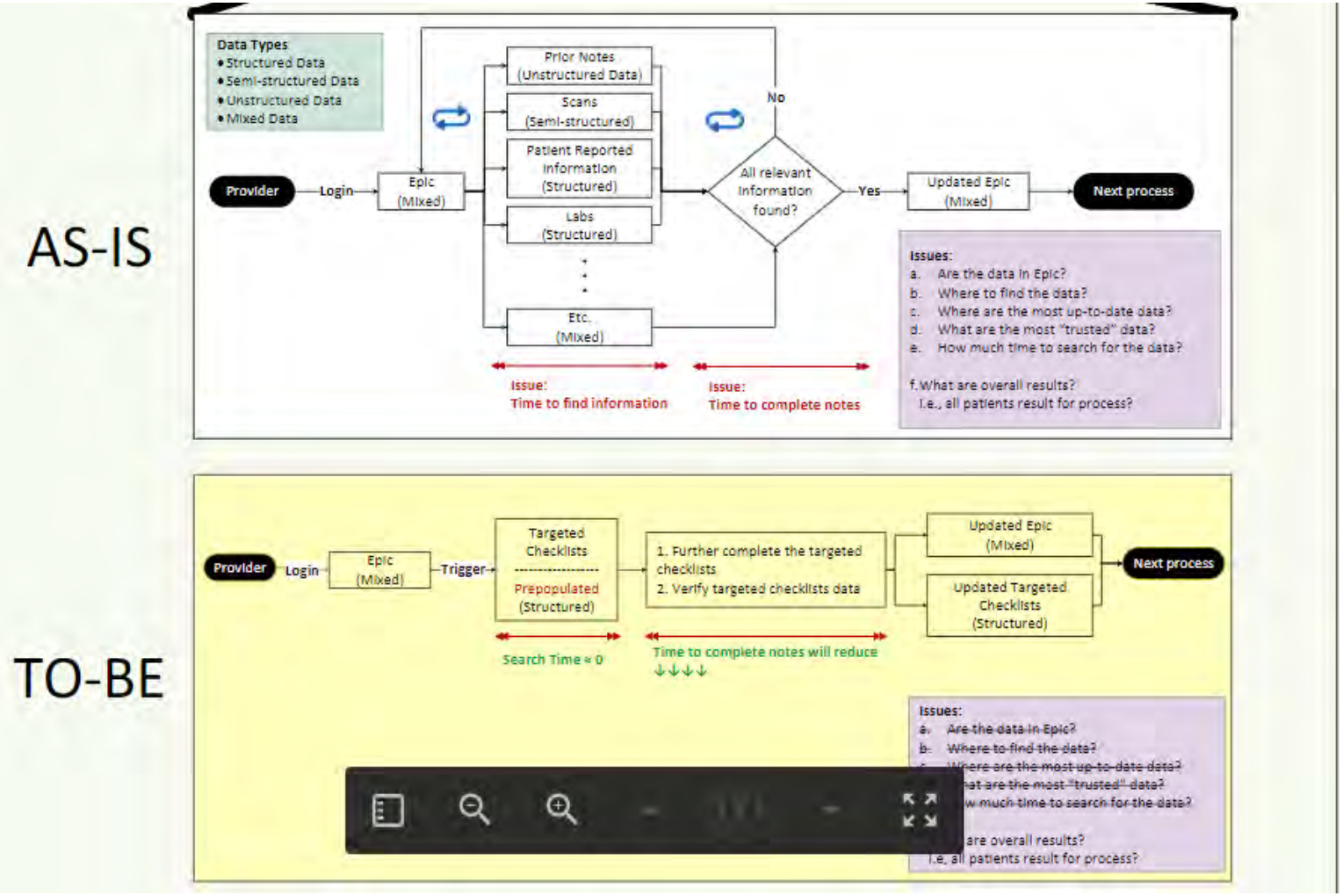
Enable improvements in technology with changes to clinical workflows



Point of Care Data Collection

Process reengineering

Starting with the AS IS and working towards the TO BE

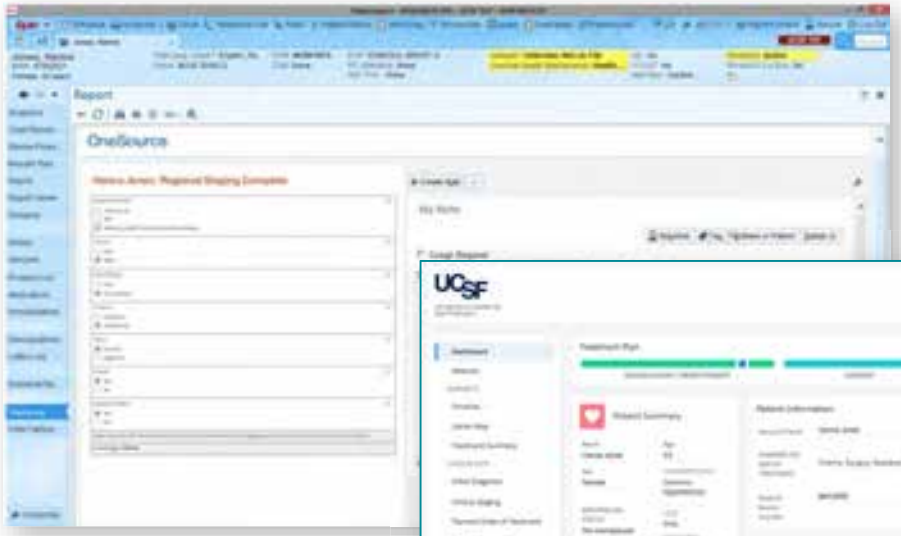


“Enter Once, Use Many”

Domain / Data Elements	New PATIENT (AT ANY PHASE)	SCREENING	DIAGNOSTIC	TREATMENT PLANNING	SURGERY	SYSTEMATIC TREATMENT	RADIATION TREATMENT	FOLLOW-UP	RESEARCH
Patient-Reported Outcomes	✎	✎	✎	✎	✎	✎	✎	✎	✎+
Patient Health History	✎	✎	✎	✎	✎	✎	✎	✎	✎+
Imaging		✎	✎	👁	👁	👁	👁	👁	✎+
Biopsy Pathology			✎	👁	👁	👁	👁	👁	✎+
Clinical Exam and Stage			✎	✎	✎	✎	✎	👁	✎+
Clinical Trial Matching			✎	✎	✎	👁	👁	👁	✎+
Treatment				✎	✎	✎	✎	👁	✎+
Final Pathology					✎	👁	👁	👁	✎+

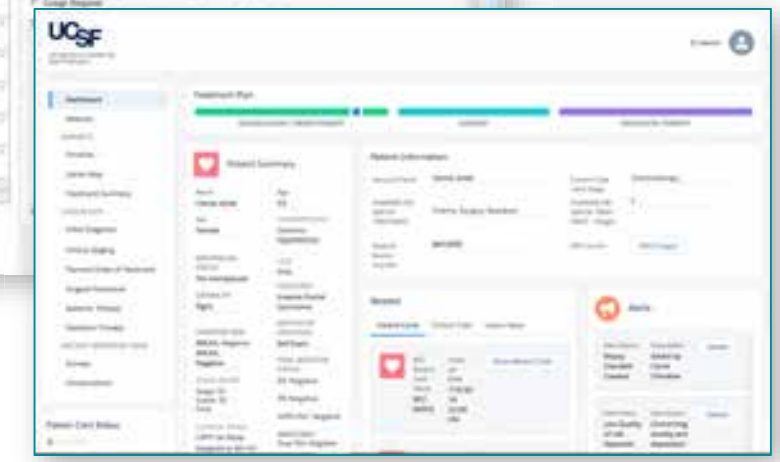
✎ New data ✎ Confirm or Update 👁 View Only ✎+ Confirm/Additional Data Added

Supporting clinical trials and data submissions

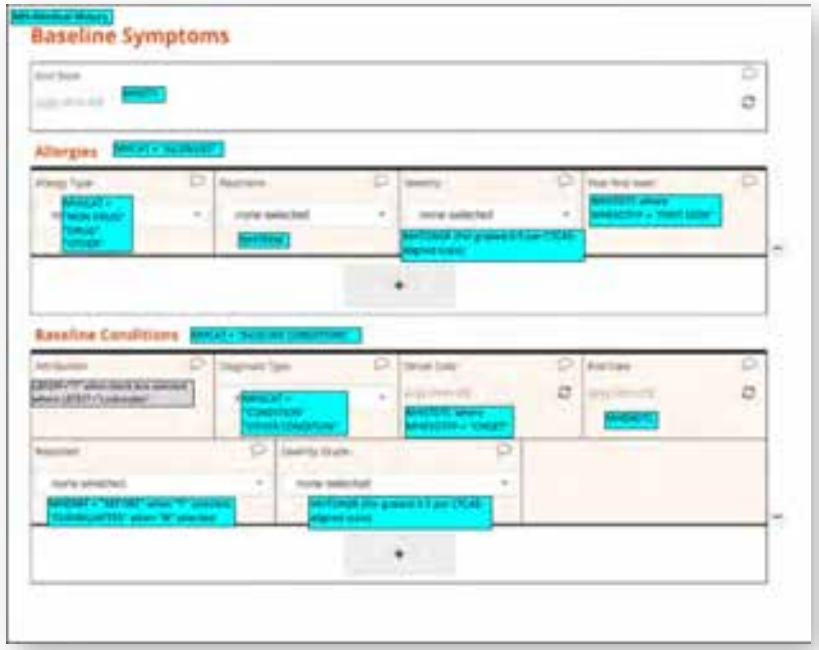


Electronic Health Record Systems

SMART on FHIR



Decision Support at Point of Care



cdisc

- TAUG-BrCa
- S(AE)
- CTCAE

Clinical Trial Data Submissions & Standards

Mobile device - Patient Reported Outcomes, Adverse Event Reporting

Session I: Establishing a High-Quality RWD Ecosystem

Session I: Establishing a High-Quality RWD Ecosystem

Wendy Rubinstein, MD, PhD
CancerLinQ / ASCO

Developing Real-World Data and Evidence to Support Regulatory Decision-Making

October 3, 2019

ASCO
AMERICAN SOCIETY OF CLINICAL ONCOLOGY

CancerLinQ is in a unique position to evaluate interoperability

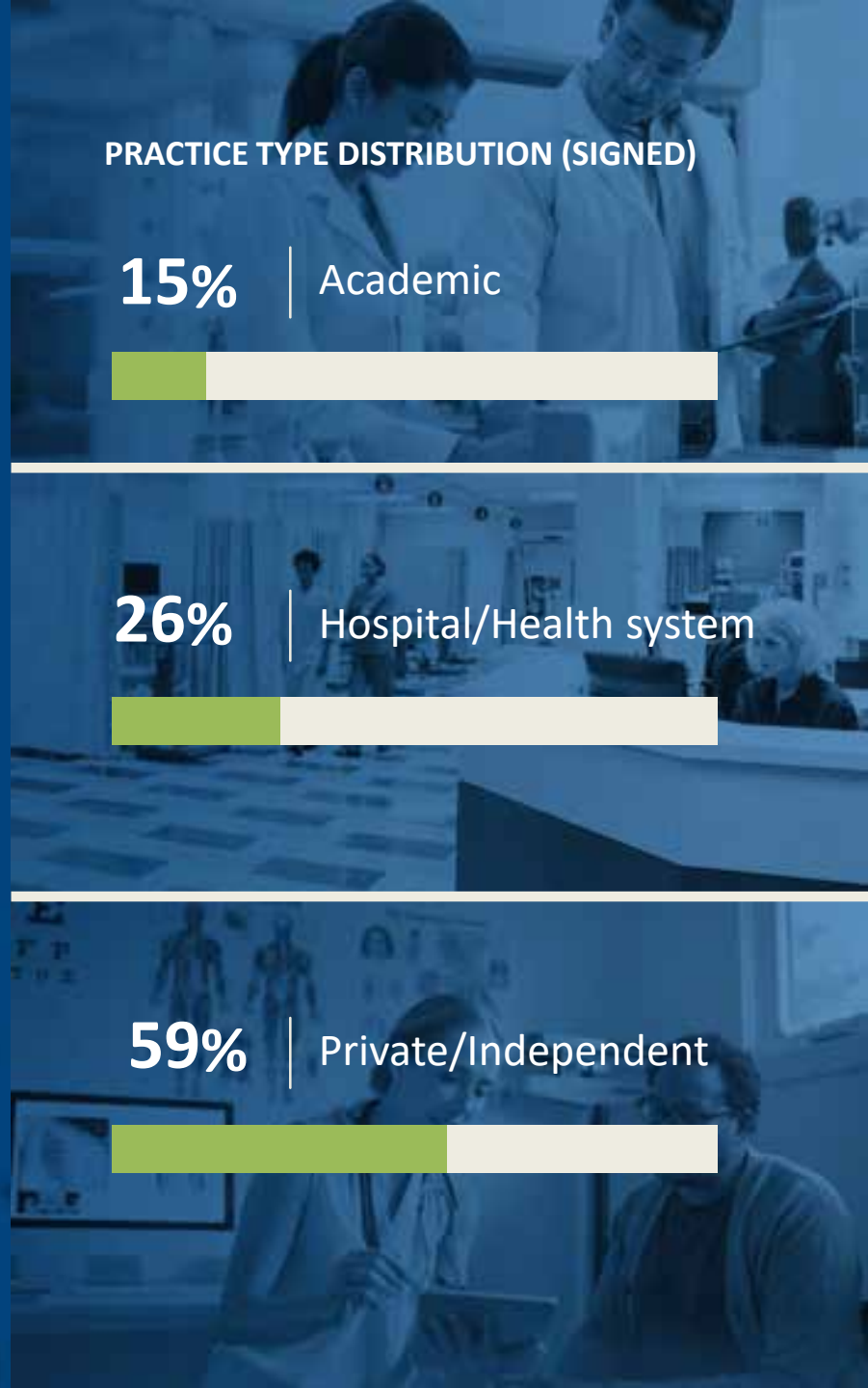
100+ | **Organizations** have signed BAAs

50+ | **Organizations** have been connected to the CancerLinQ® platform

10 | **Supported EMRs:**
Epic, MOSAIQ, Allscripts, ARIA, CureMD, OncoEMR, Integra Connect, Centricity, NextGen, IntelliDose

1,100,000+ | Total number of **patients** with a primary cancer diagnosis in the clinical database

169,000+ | **Curated records**:
Lung (NSCLC, SCLC), breast, ovarian, prostate, pancreatic, colorectal, CLL



mCODE™

Purpose: To develop and maintain standard computable data formats, known as Minimal Common Oncology Data Elements (mCODE), to achieve data interoperability and enable progress in clinical care quality initiatives, clinical research, and healthcare policy development

https://mcodeinitiative.org/

ASCO
AMERICAN SOCIETY OF CLINICAL ONCOLOGY

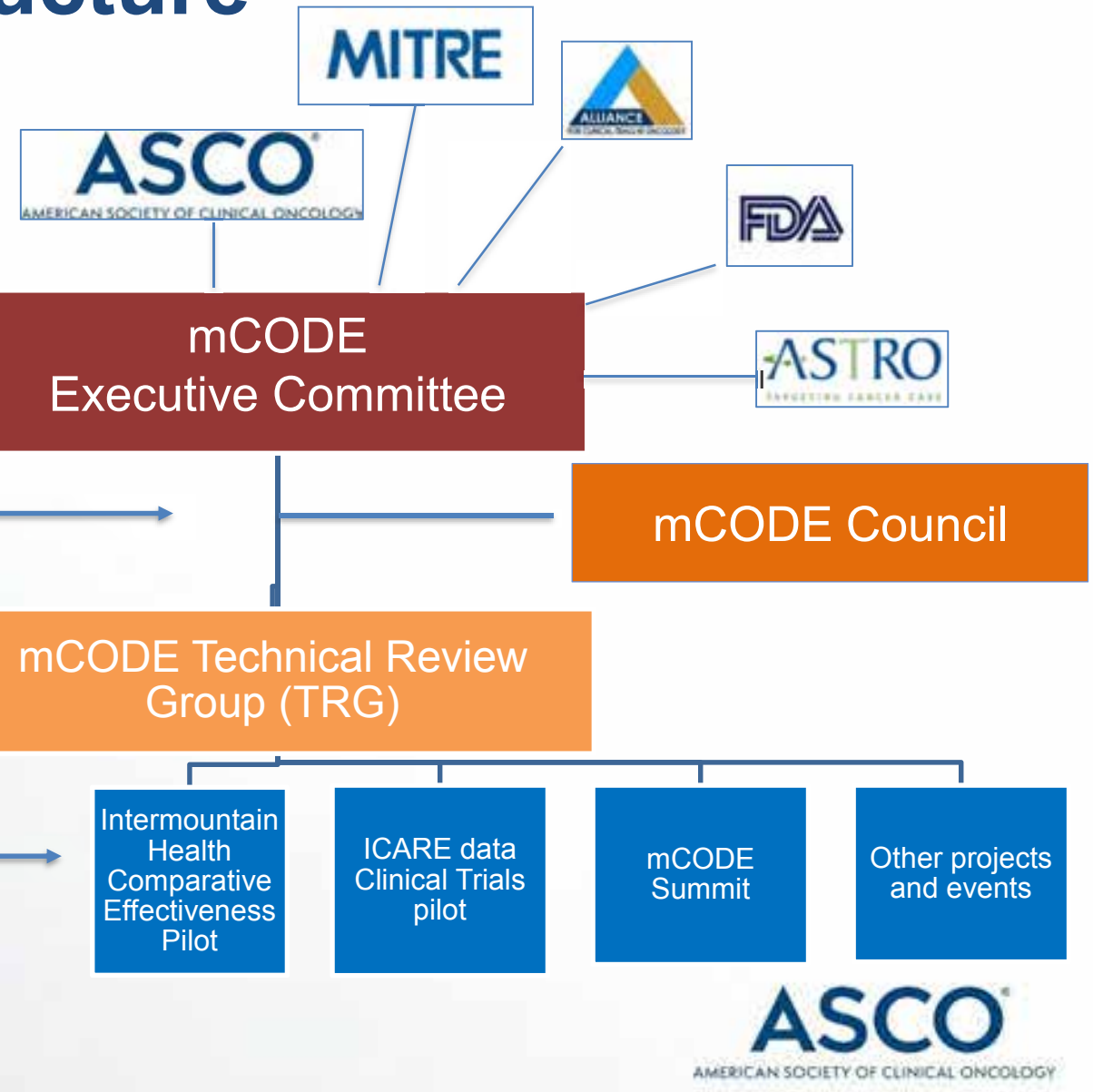
mCODE™ Governance Structure

Decision-making: Approve use cases for development, assemble and manage TRG

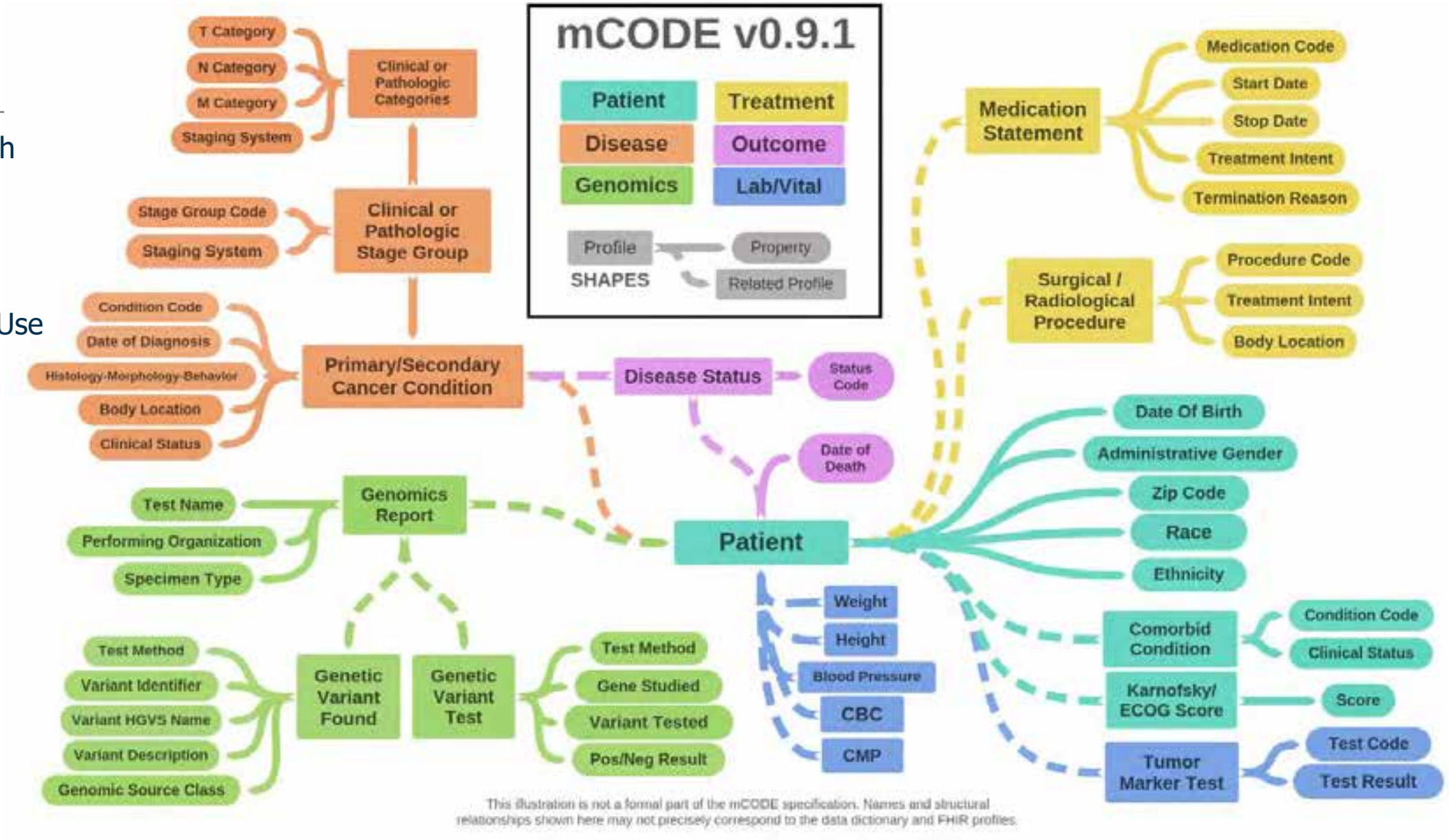
Advisors to EC: Use case sponsors, content experts, pool of potential Working Group members

Maintains mCODE data dictionary: Initial review of use cases. Convened, resourced, and managed by the EC

Working Groups: User groups assembled in response to use cases approved by EC to do the work of developing and testing new data elements



Proceeding through the HL7 balloting process for Standard for Trial Use based on FHIR R4



MITRE

Session I: Establishing a High-Quality RWD Ecosystem

Session I: Establishing a High-Quality RWD Ecosystem

Nancy Yu
CEO, RDMD

Duke-Margolis: Developing Real-World Data and Evidence to Support Regulatory Decision-Making
October 3, 2019

RDMD is a platform that helps to identify patients & generate evidence to enable drug research in rare disease

Research-activated patients

Patient recruitment & engagement

Research e-consent via central IRB protocol

RDMD obtains medical records from any facility

Longitudinal updates to records

Regulatory-ready evidence

Comparator arms based on natural history

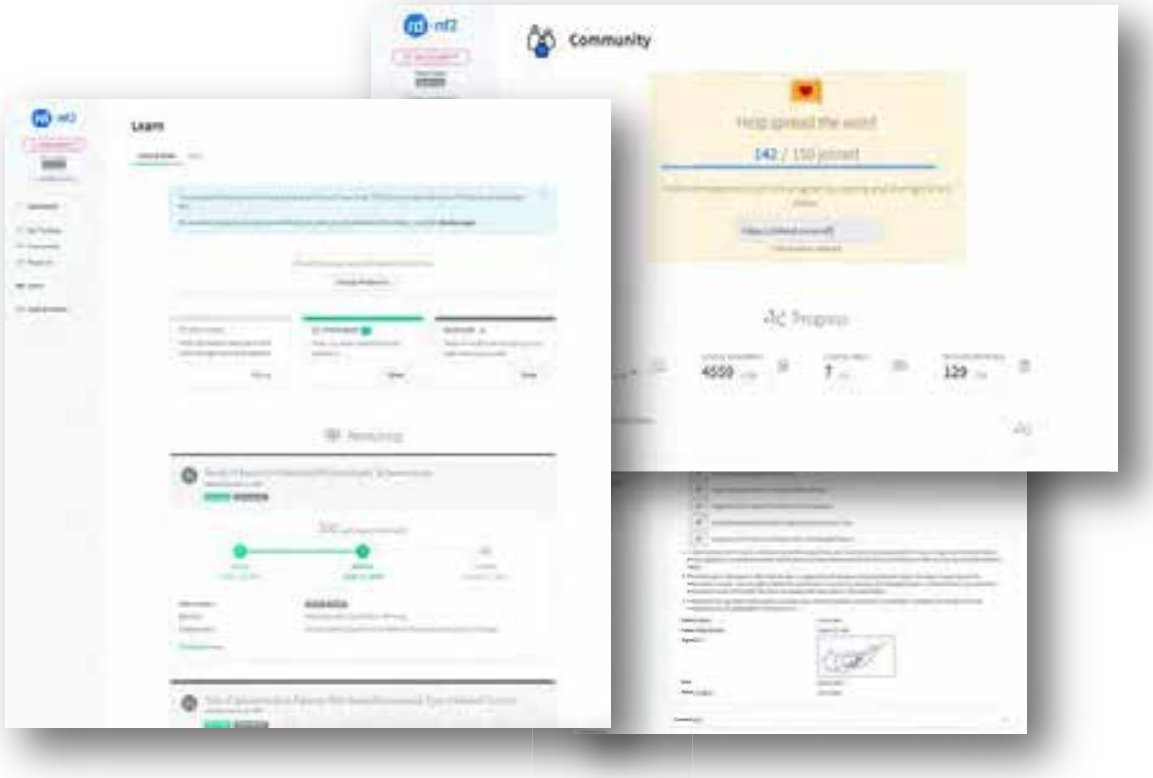
Trial qualification based on I/E criteria

Real-World Evidence on clinical outcomes

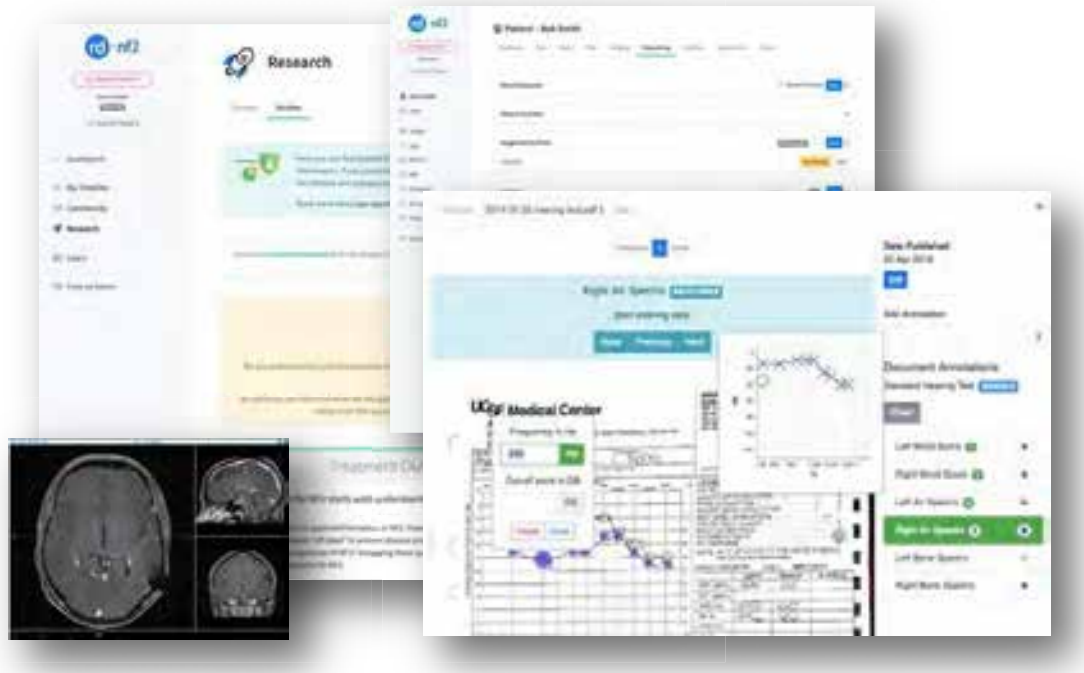
Healthcare resource utilization data

We specialize in the unique patient, community, & regulatory needs in **rare disease**

Two-sided software platform for patients & researchers

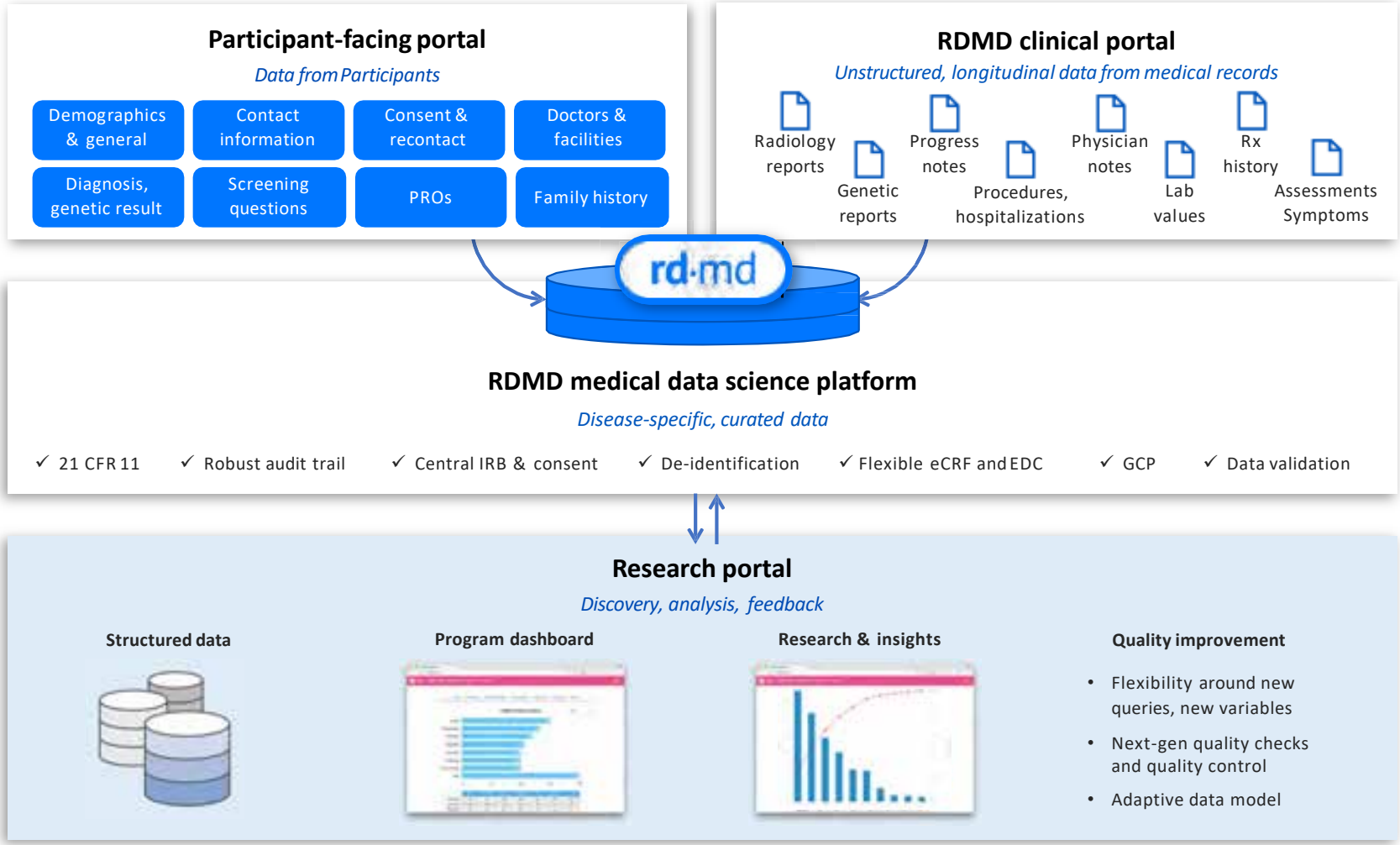


For Patients



For Life Sciences

The RDMD platform aggregates input from patients & curated data from unstructured medical records

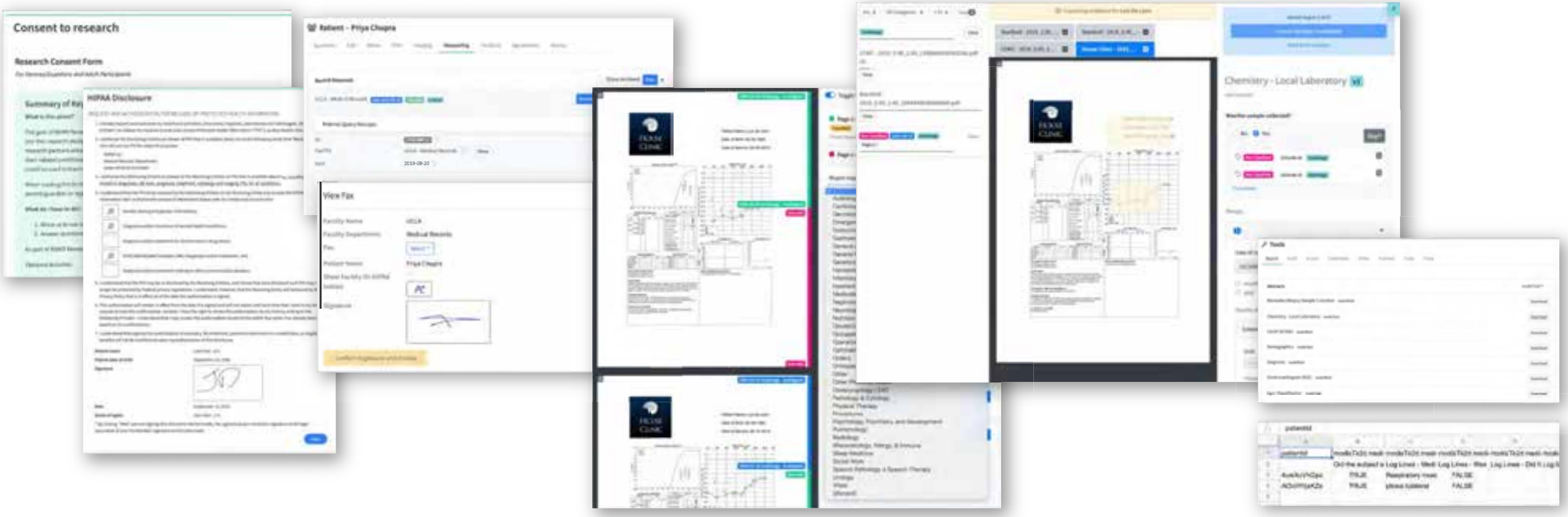


Data management challenges in rare disease

Standard of care is often poorly defined or not broadly adopted	Variability in analyzing data across different sites
Clinical outcomes assessments may 1) not be used, 2) anecdotally used, or 3) inconsistently / subjectively recorded	Incomplete data used to inform endpoint validation
Limited overall understanding of conditions to interpret complex clinical data	Difficulty in developing standard policies & procedures
Curated data is not equal to standardized data	Ensuring harmonization with existing standards is not always pragmatic
Dispersed populations requires data from disparate EHRs	Rigorous standards development & quality control needed

Our technology platform enables end-to-end **Data Quality Control** & **Data Relevancy**

Patient E-consent ▶ **Record Requesting** ▶ **Record Processing** ▶ **Record Abstraction** ▶ **Export & Analysis**



Data Quality Control: Technology, processes, training

E-signature capture

ID verification

Two-factor authentication

IRB update content management system

Hospital data verification

Automated API for e-fax & receipt

Digital & physical audit trails

Source document review (ALCOA)

QC: document verification & classification

Records attributed to Participant & processing staff

System permissions based on user roles

Continuous data quality monitoring

Data monitoring on both dataset level & abstractor level

Abstraction quality measured by quality spot audits & percentage of cases that are abstracted in duplicate

Data transfer / export checks

Periodic audits around consent & protocol scope

Data Quality Control: Data abstraction conducted under a central research protocol

Trained abstractors

Trained abstractors with clinical research or nurse practitioner backgrounds

RDMD technology platform

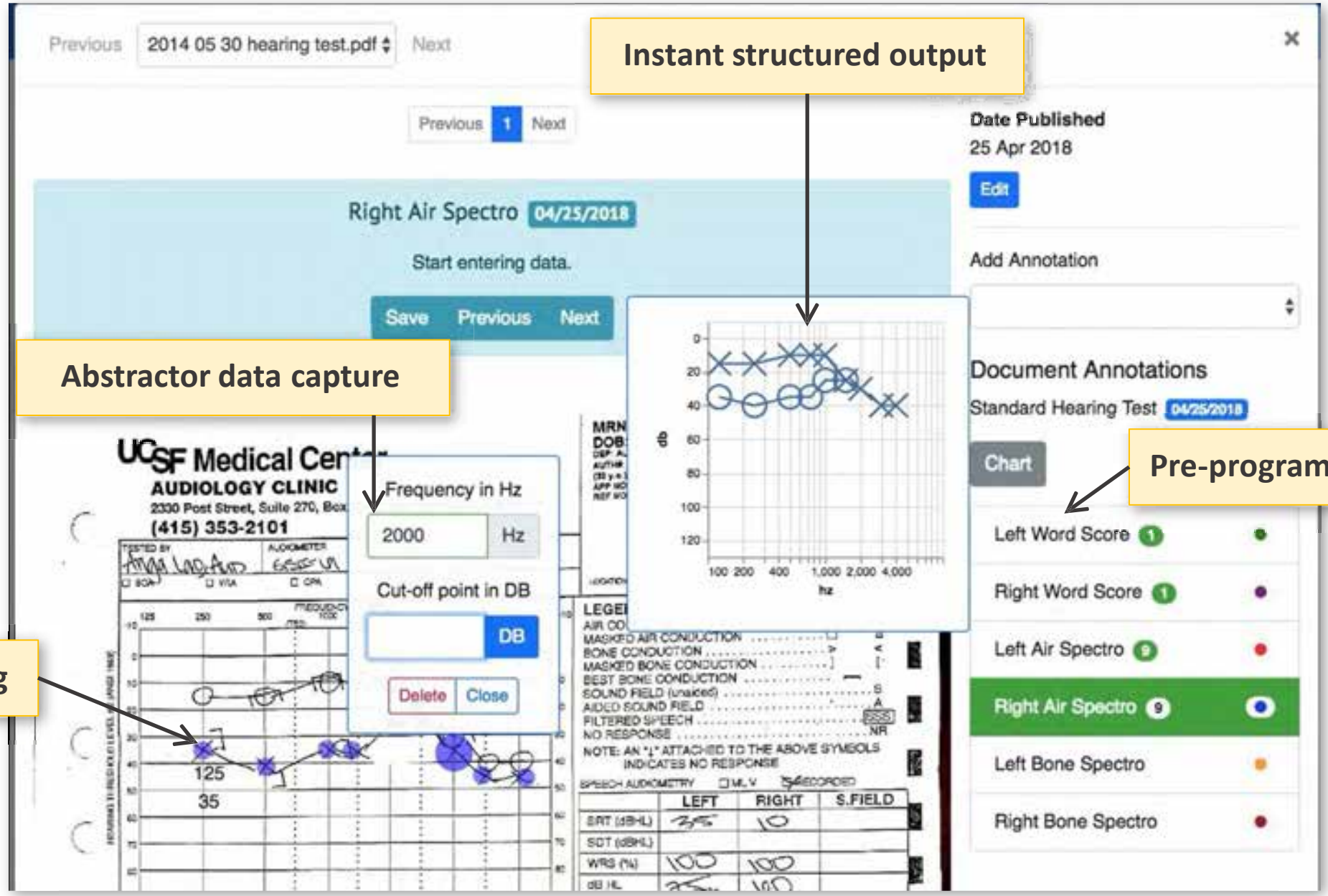
Software enables effective document review & data capture in predetermined forms

RDMD centralized research protocol

Umbrella research protocol & patient informed consent form allows for:

- Flexible / adaptive data capture protocols
- Broad research use on de-identified prospective & retrospective data
- Patient recontact for future studies
- Data analysis across diseases

rd·md


Previous
2014 05 30 hearing test.pdf
Next
Previous 1 Next
Instant structured output
Date Published
25 Apr 2018
Edit
Right Air Spectro 04/25/2018
Start entering data.
Save Previous Next
Add Annotation
Abstractor data capture
Document Annotations
Standard Hearing Test 04/25/2018
Chart
Pre-programmed forms
Frequency in Hz
2000
Hz
Cut-off point in DB
DB
Delete Close
Left Word Score 1
Right Word Score 1
Left Air Spectro 9
Right Air Spectro 9
Left Bone Spectro
Right Bone Spectro
Evidence linking
UCSF Medical Center
AUDIOLOGY CLINIC
2330 Post Street, Suite 270, Box
(415) 353-2101
125
35
LEFT
RIGHT
S.FIELD

Data Relevancy: Growing clinical module library maps to industry standards

Standard Modules

Examples:

- Diagnosis
- Demographics
- Assistive devices
- Lab values
- Echocardiograms
- Comorbidities
- Concomitant medications
- Audiology assessments
- Genetic testing
- EKGs
- Healthcare utilization

Therapeutic-Area Specific Modules

Examples:

- Clinical milestone modules
- Disease-specific symptom modules
- Disease-specific assessments
  - Urine GAG testing, MPS enzyme testing

Maps to / Conforms with

- MedDRA, WHODrug, CDISC, NINDS, GRDR, SNOMED, relevant trial protocols, literature

Data Relevancy: High confidence in real-world data requires triangulation of multiple data sources

	High Confidence	Medium Confidence	Low Confidence
	Original source documents available	Physician confirms endpoints in note, but source documents unavailable	Endpoints briefly referenced in physician note; source documents unavailable
Patient Case	Otitis media & hearing loss documented in **44 ENT notes, 9 audiograms, & referenced extensively** in physician notes	Urine GAG results **copied into note** but original report unavailable	Physician noted that patient had a "sleep study available for review showing AHI obstructive of 5," **but the study was not referenced again** & polysomnography report unavailable
	Ideal; include data: Tag all mentions of the variable to allow for a robust audit trail	**Acceptable; include data:** Tag all mentions of variable; contact patient / institution to track down source if needed	**Likely unacceptable; flagged:** Patient may be contacted to confirm all institutions

rd·md

Data Relevancy: Patients are key partners in data quality & completeness

Minimize missing data

Patients respond with key information, verifications, & critical documents

Recontact for follow-ups & future studies

Patients are informed about future research opportunities

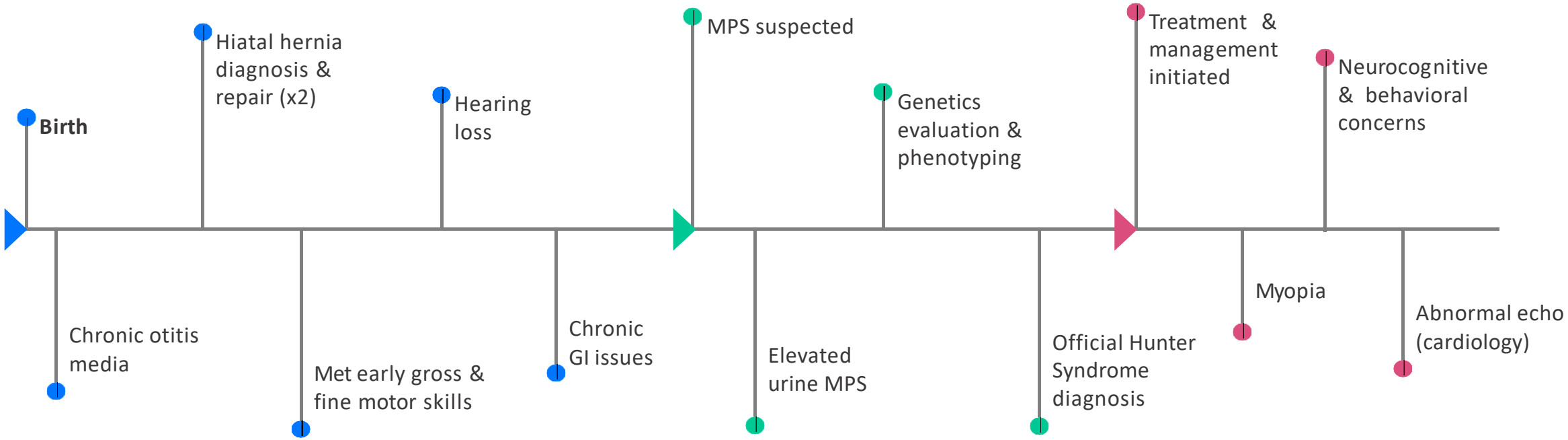
FDA: Rare Diseases: Natural History Studies for Drug Development Guidance for Industry

"Patients' continuing study participation ensures the robustness of follow-up data"

Appendix: Participant case study

Understanding early natural history, diagnosis, & management outcomes in rare is **complex,** requiring analysis of multiple sources of clinical data

Hunter Syndrome (MPS II) Case Study



What sources of data can we use to build out the typical patient journey in a rare condition?

Analysis of real-world data, *in context*, can reveal the earliest signs of rare conditions

Birth
Normal newborn screen, ABR

Newborn Screening Results

Date received: 2/XX/2012
Date Reported: 2/XX/2012

SEX: Male
Feed: Breastmilk only
Status: NORMAL
NORMAL SCREEN

Disorder	Screening Result
Amino Acid Disorders	Normal
Fatty Acid Disorders	Normal
Organic Acid Disorders	Normal
Galactosemia	Normal
Biotinidase Deficiency	Normal
Hypothyroidism	Normal
CAH	Normal
Hemoglobinopathies	Normal
Cystic Fibrosis	Normal

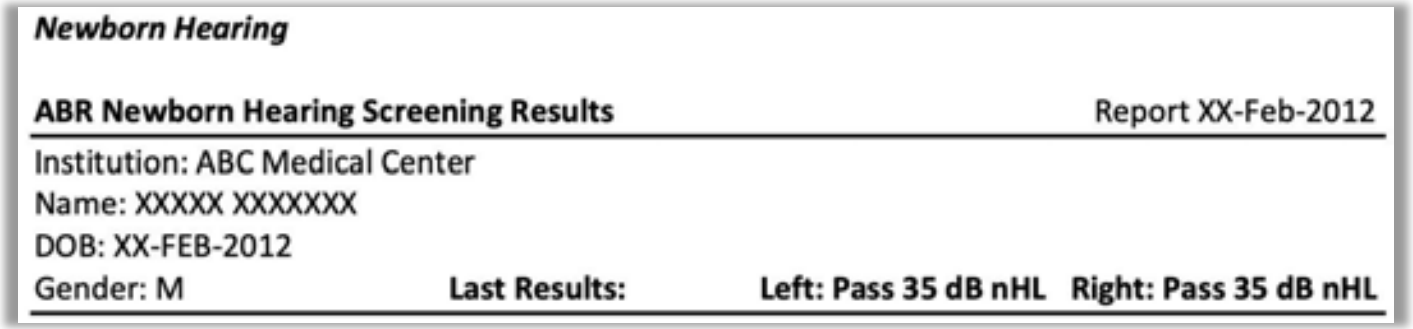
5 months:
Chronic otitis media first noted; first PE tubes placed

0-2 years:
Met early gross & fine motor skills on time

Age 3
Third set of PE tubes; intra-operative ABR identifies bilateral sensorineural hearing loss
Receives hearing aids

Age 4
Present to GI clinic for chronic diarrhea; X-ray ordered

Age 4:
Signs of MPS discovered incidentally



Newborn Hearing

ABR Newborn Hearing Screening Results Report XX-Feb-2012

Institution: ABC Medical Center
Name: XXXXX XXXXXXX
DOB: XX-FEB-2012
Gender: M **Last Results:** **Left: Pass 35 dB nHL** **Right: Pass 35 dB nHL**

Analysis of real-world data, *in context*, can reveal the earliest signs of rare conditions

ENT Notes: 44
Audiograms: 9

Birth
Normal newborn screen, ABR

5 months:
Chronic otitis media first noted; first PE tubes placed

4-17 months:
Hiatal hernia identified; repaired surgically 2X

0-2 years:
Met early gross & fine motor skills on time

Age 3
Third set of PE tubes; intra-operative ABR identifies bilateral sensorineural hearing loss
Receives hearing aids

Age 4
Present to GI clinic for chronic diarrhea; X-ray ordered

Age 4:
Signs of MPS discovered incidentally

Provider: Dr. X (Otolaryngology)

Reason for Appointment
1. Ear infection

History of Present Illness
10/XX/12: New patient in consultation for ear infections. Location: bilateral. Quality: has had 3-5 ear infections. Severity: 3/5, moderate. Duration: for 2 months. Timing: continuous. Context: Full term delivery, passed NBS, pets in house, smokers outside house, attends daycare, breastfed. Modifying factors: recurrent following antibiotics. Assoc signs / symptoms: fever with infections, has had nasal congestion and rhinorrhea, has had cough, snoring.

ENT Documentation of PE Tubes

Provider: Dr. X (Otolaryngology)

Reason for Appointment
1. Otitis media

History of Present Illness
Postop: Post-op appt for BMT done on 10/XX/12 at ABC Hospital. Mom says she has noticed the improvement with this hearing.
Surgical Procedure: Myringotomy with PE tube. Location: Bilateral. Surgery Date: 10/XX/12.

Pre-Diagnosis

Analysis of real-world data, *in context*, can reveal the earliest signs of rare conditions

RADIOLOGY REPORT
ESOPHAGRAM

Attending Physician: Dr. Z (Radiologist)

DATE: 6/XX/2012
INDICATION: Stridor. Assess for vascular ring.

IMPRESSION:
1. Normal esophagus with no findings of a vascular ring
2. Sliding or paraesophageal hiatal hernia

Birth
Normal newborn screen, ABR

5 months:
Chronic otitis media first noted; first PE tubes placed

4-17 months:
Hiatal hernia identified; repaired surgically 2X

0-2 years:
Met early gross & fine motor skills on time

Age 3
Third set of PE tubes; intra-operative ABR identifies bilateral sensorineural hearing loss
Receives hearing aids

Age 4
Present to GI clinic for chronic diarrhea; X-ray ordered

Age 4:
Signs of MPS discovered incidentally

DISCHARGE SUMMARY

Admitting Physician: Dr. Y
Date of Admission: 11/XX/12
Date of Discharge: 11/XX/12
Admission Diagnosis: Hiatal hernia and GERD
Discharge Diagnosis: Hiatal hernia. GERD s/p Nissen Fundoplication
Operative Procedures: Laparoscopic hiatal hernia repair and Nissen fundoplication

SURGICAL PATHOLOGY REPORT

Attending Physician: Dr. A (Pathologist)

Collected Date: 7/XX/2013

Diagnosis
Soft tissue, hiatal hernia, excision: consistent with hernia with focal acute serositis and reactive serosal changes

rd·md

Pre-Diagnosis

Analysis of real-world data, *in context*, can reveal the earliest signs of rare conditions

RIGHT EAR – testing was performed by Dr. B. AuD, with ABC ENT Specialists.
Sedated ABR predicts a mild sensorineural hearing loss.

Right Ear	500Hz	1000Hz	2000Hz	4000Hz
Airdr (dBHL)	25	40	40	30

LEFT EAR – testing was performed by Dr. B. AuD, with ABC ENT Specialists.
Sedated ABR predicts a mild to moderate sensorineural hearing loss.

Right Ear	500Hz	1000Hz	2000Hz	4000Hz
Airdr (dBHL)	40	45	45	35

Birth
Normal newborn screen, ABR

5 months:
Chronic otitis media first noted; first PE tubes placed

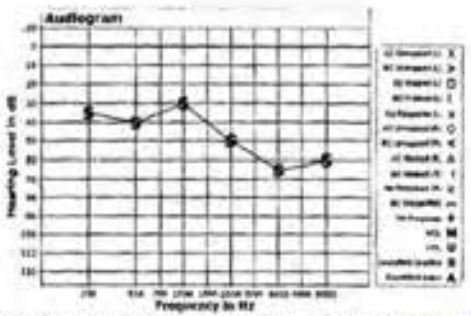
4-17 months:
Hiatal hernia identified; repaired surgically 2X

Age 3
Third set of PE tubes; intra-operative ABR identifies bilateral sensorineural hearing loss
Receives hearing aids

Age 4
Present to GI clinic for chronic diarrhea; X-ray ordered

Age 4:
Signs of MPS discovered incidentally

Audiogram
Audiologist: Dr. C., Au.D.
Visit Date: 3/XX/2015



Results: Threshold Re-Check: Patient was tested via play audiometry in the sound field by Dr. C., Au.D. Responses to warble tone stimuli indicate a mild to moderate sensitivity loss in at least the better ear.

Analysis of real-world data, *in context*, can reveal the earliest signs of rare conditions

Birth
Normal newborn screen, ABR

5 months:
Chronic otitis medi
first noted; first PE
tubes placed

XR Abdomen 1
Date: 11/XX/2016
Resulted by: Dr. K (Radiologist)

CLINICAL HISTORY:
Reason for Exam: Evaluate for constipation
Clinical Signs and Symptoms: Loose stools with suspected constipation

COMPARISON: None

FINDINGS
Catheters / tubes / postoperative changes: None
Bowel: Normal bowel gas pattern
Soft tissues: The liver appears prominent, and clinical correlation is advised. The configuration of the pelvis and the ribs raises the possibility of an underlying storage disease such as mucopolysaccharidosis.
Lung bases: Clear
Bones: Lack of complete posterior spinal fusion at the L5 level is a finding of uncertain clinical significance.

IMPRESSION: No evidence of excessive stool burden. The bones raise the possibility of an underlying storage disease. Probable hepatomegaly.

Age 4
Present to GI clinic for chronic diarrhea; X-ray ordered

Age 4:
Signs of MPS discovered incidentally

Clinical, radiological, and laboratory data can help answer key questions about a patient's diagnostic odyssey

Age 4:
Signs of MPS discovered incidentally

Genetics Clinic Note

Provider: Dr. P (Genetics)
Encounter Date: 12/XX/2016
Chief Complaint: Possible mucopolysaccharidosis
HPI: ….He was most recently evaluated by Dr. J in Gastroenterology for diarrhea and chronic loose stools. An abdominal x-ray did not reveal an underlying etiology, but noted hepatomegaly and bone changes that could represent a storage disorder. We had recommended a urine MPS screen before seeing him in Genetics clinic, which was recently resulted.

Age 4:
Urine glycosaminoglycans ordered

Age 4:
Hunter diagnosis confirmed

Management begins

Quantitative Urine MPS (External Lab X)

Mucopolysaccharides (MPS) <=16.0 mg/mmol	**98.6**	**H**

Note: Lab results sourced from genetics clinic note

Clinical, radiological, and laboratory data can help answer key questions about a patient's diagnostic odyssey

Age 4:
Signs of MPS discovered incidentally

Age 4:
Initial Genetics evaluation identifies additional signs of Hunter syndrome

Age 4:
Urine glycosaminoglycans ordered

Genetics Clinic Note

Provider: Dr. P (Genetics)
Encounter Date: 12/XX/2016
Chief Complaint: Possible mucopolysaccharidosis
Physical Examination:

General: Awake, alert, no acute distress.
HEENT: **Coarse facial appearance with midface hypoplasia, deep philtrum, broad eyebrows.** Pupils are equal and reactive to light. No corneal clouding appreciated. Mouth normal, dentition and oropharynx clear.
Neck: Supple
Chest: Normal in shape and configuration. Clear to auscultation bilaterally.
CVS: Regular in rate and rhythm. No murmurs, rubs or gallops appreciated on auscultation.
Abdomen: Soft, non-tender. **Mild hepatomegaly, palpable about 2-3 cm below the costal margin.** No splenomegaly appreciated. **Mild umbilical hernia.**
GU: Normal male genitalia
Back: Intact
Extremities: **Flexion contractures of the hands. Tapered, almost trident appearance of fingers and hands.** All digits intact. Warm and well perfused, with brisk capillary refill. **Decreased range of motion shoulders.**
Neurological Exam: Cranial nerves appear grossly intact. There is normal muscular strength and [illegible]re normal and symmetric. Sensation appears intact. No cerebellar [illegible]l movements noted on examination.

Impression: The patient is a 4-year-old male with a history of **speech delay, recurrent otitis media, bilateral sensorineural hearing loss, hepatosplenomegaly, coarse facial features, and urine mucopolysaccharides and x-ray results that are consistent with a lysosomal storage disorder, most likely a form of mucopolysaccharidosis (MPS).**

Based on his features and medical history, we feel that he most likely has MPS I (Hurler, Hurler-Scheie) or MPS II (Hunter Syndrome), both of which have ERT available. However, we cannot determine which one he has without additional enzymatic testing.

Clinical, radiological, and laboratory data can help answer key questions about a patient's diagnostic odyssey

Age 4:
Signs of MPS discovered incidentally

Age 4:
Initial Genetics evaluation identifies additional signs of Hunter syndrome

Diagnostic Testing
Type of Specimen: Leukocytes
Test Requested: MPS Panel
Sample Collection Date: 12/XX/2016
Sample Received: 12/XX/2016
Date Reported: 1/XX/2017

	Results	Normal	Note
Alpha-iduronidase	23.66	6-714	Normal
Iduronate 2 sulfatase	<3.95	155-1082	BQL
Heparan N-sulfatase (type A)	10.5	2.21-27.34	Normal
N acetyl alpha glucosaminidase (type B)	393.7	97-1064	Normal
Acetyl CoA: glucosamine N acetyl transferase (type C)	13.7	5.8-45	Normal
N acetyl glucosamine 6 sulfatase (type D)	15.6	5.3-35.4	Normal
N acetyl galactosamine 6 sulfatase	212.3	49-255	Normal
Beta galactosidase	81.42	13.5-176	Normal
Arylsulfatase B	89.6	57-493	Normal
Beta glucuronidase	103.51	31.4-224	Normal

Management begins

Psychology Clinic Note

Provider: Dr. T (Psychology)
Encounter Date: 2/XX/2017
Reason for Referral: He is a 4-year-old boy. He was referred for evaluation by his geneticist due to concerns of neurocognitive dysfunction associated with his medical condition. He was recently diagnosed with a metabolic storage disorder (Hunter Syndrome).

Age 4:
Hunter diagnosis confirmed

Metabolic Genetics Note

Provider: Dr. M (Metabolic Genetics)
Encounter Date: 7/XX/2017
HPI: He is a 5-year-old male with a recent diagnosis of MPS type II (Hunter Syndrome).

Note: Lab results sourced from the 7/2017 metabolic genetics note

Analysis of the post-diagnostic journey allows for tracking of long-term outcomes

Age 4:
Initiation of IV idursulfase

Age 4:
Myopia diagnosed; no retinopathy

Age 6:
Bicuspid aortic valve, thickened mitral valve noted on echocardiogram

Age 6:
Ongoing neurocognitive concerns; autism diagnosis

Provider: Dr. X (Otolaryngology)
Encounter Date: 2/X/2017

Current Medications
Taking multivitamin Multiple Vitamins tablet, chewable, 1 tab(s) once a day
Taking Elaprase 2mg/mL solution once a week
Taking Zyrtec 1mg/mL syrup 10mL once a day

Analysis of the post-diagnostic journey allows for tracking of long-term outcomes

Age 4:
Initiation of IV idursulfase

Ongoing
Monitoring via ENT, ophthalmology, cardiology, GI, orthopedics, developmental pediatrics, genetics

Age 4:
Myopia diagnosed; no retinopathy

Age 6:
Bicuspid aortic valve, thickened mitral valve noted on echocardiogram

OPHTHALMOLOGY CLINIC NOTE

Provider: Dr. O (Ophthalmology)
Encounter Date: 12/X/2016
Impression: Likely MPS; no corneal clouding noted on limited anterior segment examination (seen usually with MPS I and MPS VI); no evidence of RPE changes to suggest retinopathy (often seen with MPS I, MPS II, and MPS III); optic nerve appearance today with wnl, though changes can develop over time; pt would not allow IOP check, but globes felt soft to palpation OU (reports of glaucoma in MPS I, II, VI). In addition, amblyopia, strabismus, and hyperopia have been described in MPS (of note, patient has myopia, while hyperopia is more commonly described in MPS I, and astigmatism, which has been seen in MPV IV).

Analysis of the post-diagnostic journey allows for tracking of long-term outcomes

Age 4:
Initiation of IV idursulfase

Assessment performed by: Dr. T (Neuropsychology)
Date of Assessment: February 2017
Wechsler Preschool and Primary Scale of Intelligence- 4th Edition

	SS	%ile
Full Scale IQ	101	53
Verbal Comprehension Composite	99	47
Visual Spatial Composite	97	42
Fluid Reasoning Composite	100	50

Age 4:
Myopia diagnosed; no retinopathy

Assessment performed by: Dr. T (Neuropsychology)
Date of Assessment: February 2018
Wechsler Preschool and Primary Scale of Intelligence- 4th Edition

	SS	%ile
Full Scale IQ	80	9
Verbal Comprehension Composite	81	10
Visual Spatial Composite	78	7
Fluid Reasoning Composite	85	16

NEUROPSYCHOLOGICAL EVALUATION CLINIC NOTE

Provider: Dr. T (Psychology)
Encounter Date: 2/X/2017
Impression: The patient's current neuropsychological profile reflects solidly average intelligence and memory along with weakness in visual-motor integration and significant symptoms of hyperactivity / impulsivity and aggressiveness. To a lesser degree there also is some evidence of challenges with attentional control and moodiness. This pattern of performance appears to be consistent with expectation for Hunter Syndrome.

Diagnoses:
Neurocognitive dysfunction seconda
ADHD, predominantly hyperactivity /

Age 4:
Neurocognitive & behavioral concerns noted; ADHD diagnosis

Age 6:
Ongoing neurocognitive concerns; autism diagnosis

NEUROPSYCHOLOGICAL EVALUATION CLINIC NOTE

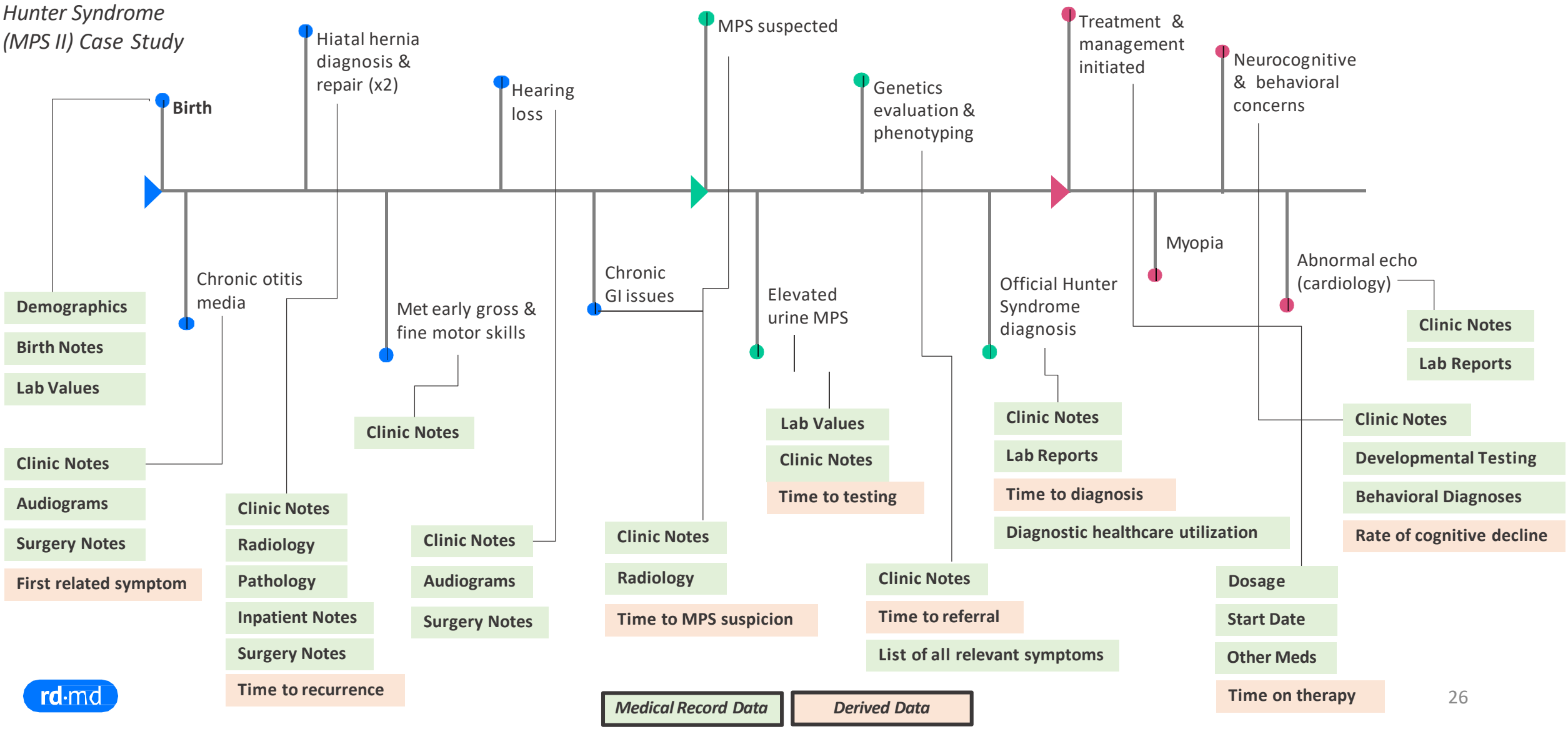
Provider: Dr. T (Neuropsychology)
Encounter Date: 2/X/2018
Impression: Compared to prior results, there appears to have been slowing in developmental progression across many domains, including intellectual capacity, memory, and visual-motor integration. However, it is important to be mindful of his degree of attentional dysregulation, which could be impacting his performance in these domains.

In sum, there appears to have been some slowing in overall neurocognitive development, although in part that could be attributable to increased attentional dysregulation. It is unlikely that there has been deterioration in cognitive reasoning or memory at this point in time.

BEHAVIORAL EVALUATION
DOCUMENTED PARENT LETTER

Provider: Dr. P (Behavioral Psychology at External Institution)
Encounter Date: 2/X/2019
Summary: The patient meets the DSM-5 symptom criteria for an Autism Spectrum Disorder (ASD). He exhibits developmental difficulty in the areas of both social communication and non-verbal communication. His social communication is far below that expected for his general developmental delay.

Evaluation of multiple source documents is required to understand the patient journey in rare disease



Thank you!

Nancy Yu, *CEO* nancy@rdmd.com
Kristina Cotter, PhD, CGC, MS, *Research Director* kristina@ rdmd.com

rd·md

Session I: Establishing a High-Quality RWD Ecosystem

Collaborating Organizations

Lead – HPHC Institute

Data & Scientific Partners

Scientific Partners

https://www.sentinelinitiative.org/collaborators

Available Data Elements

Administrative Data

Enrollment	Demographic	Dispensing	Encounter	Diagnosis	Procedure
Patient ID	Patient ID	Patient ID	Patient ID	Patient ID	Patient ID
Enrollment Start & End Dates	Birth Date	Dispensing Date	Service Date(s)	Service Date(s)	Service Date(s)
Drug Coverage	Sex	National Drug Code (NDC)	Encounter ID	Encounter ID	Encounter ID
Medical Coverage	Zip Code	Days Supply	Encounter Type and Provider	Encounter Type and Provider	Encounter Type and Provider
Medical Record Availability	Etc.	Amount Dispensed	Facility	Diagnosis Code & Type	Procedure Code & Type
			Etc.	Principal Discharge Diagnosis	Etc.

Clinical Data

Lab Result	Vital Signs
Patient ID	Patient ID
Result & Specimen Collection Dates	Measurement Date & Time
Test Type, Immediacy & Location	Height & Weight
Logical Observation Identifiers Names and Codes (LOINC®)	Diastolic & Systolic BP
Etc.	Tobacco Use & Type
	Etc.

Registry Data

Death	Cause of Death	State Vaccine
Patient ID	Patient ID	Patient ID
Death Date	Cause of Death	Vaccination Date
Source	Source	Admission Date
Confidence	Confidence	Vaccine Code & Type
Etc.	Etc.	Provider
		Etc.

Inpatient Data

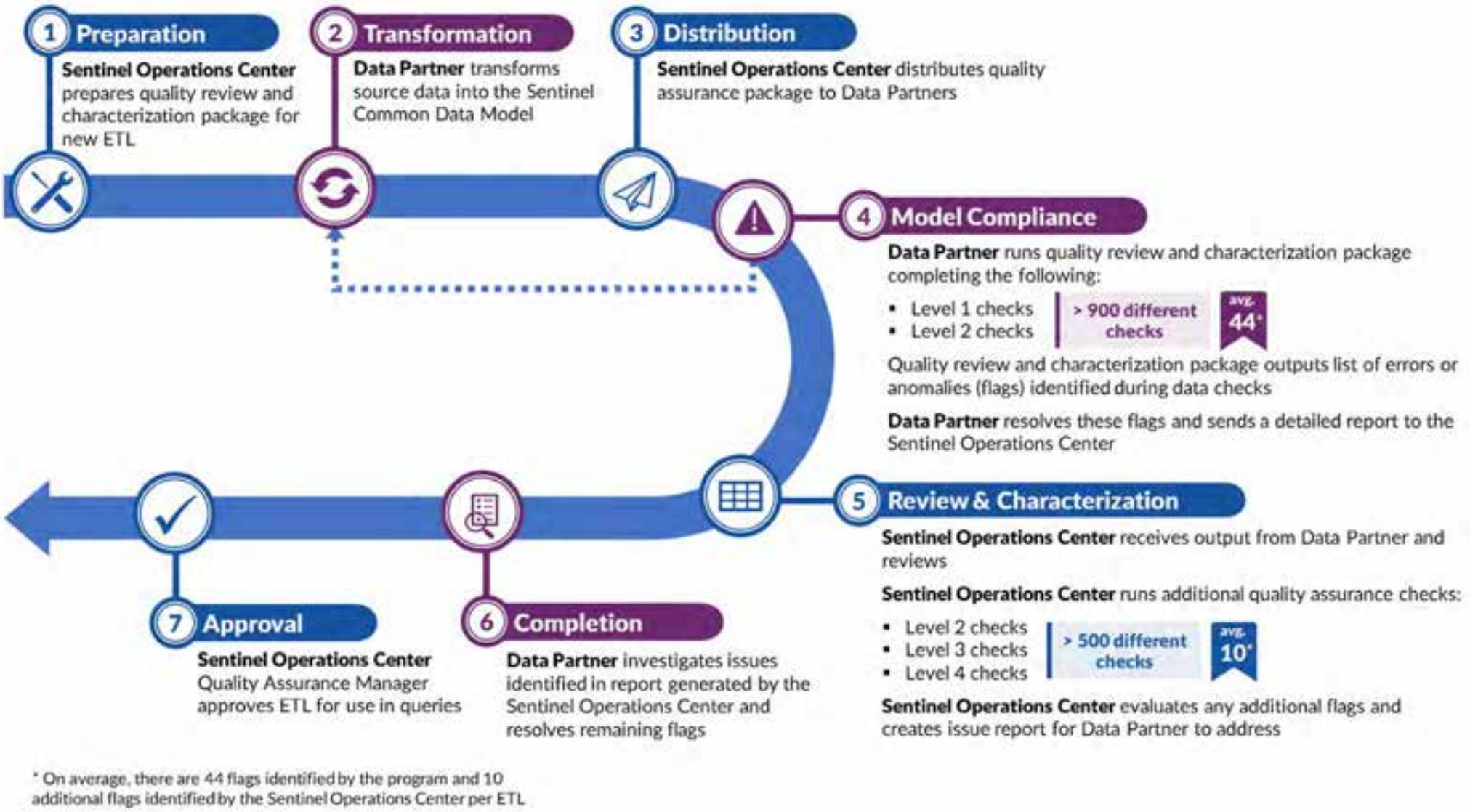
Inpatient Pharmacy	Inpatient Transfusion
Patient ID	Patient ID
Administration Date & Time	Administration Start & End Date & Time
Encounter ID	Encounter ID
National Drug Code (NDC)	Transfusion Administration ID
Route	Transfusion Product Code
Dose	Blood Type
Etc.	Etc.

Mother-Infant Linkage Data

Mother-Infant Linkage
Mother ID
Mother Birth Date
Encounter ID & Type
Admission & Discharge Date
Child ID
Child Birth Date
Mother-Infant Match Method
Etc.

https://www.sentinelinitiative.org/sentinel/data/distributed-database-common-data-model

Data Quality Review and Characterization Process



https://www.sentinelinitiative.org/sentinel/data-quality-review-and-characterization

Data Quality Checks and Examples

https://www.sentinelinitiative.org/sentinel/data-quality-review-and-characterization

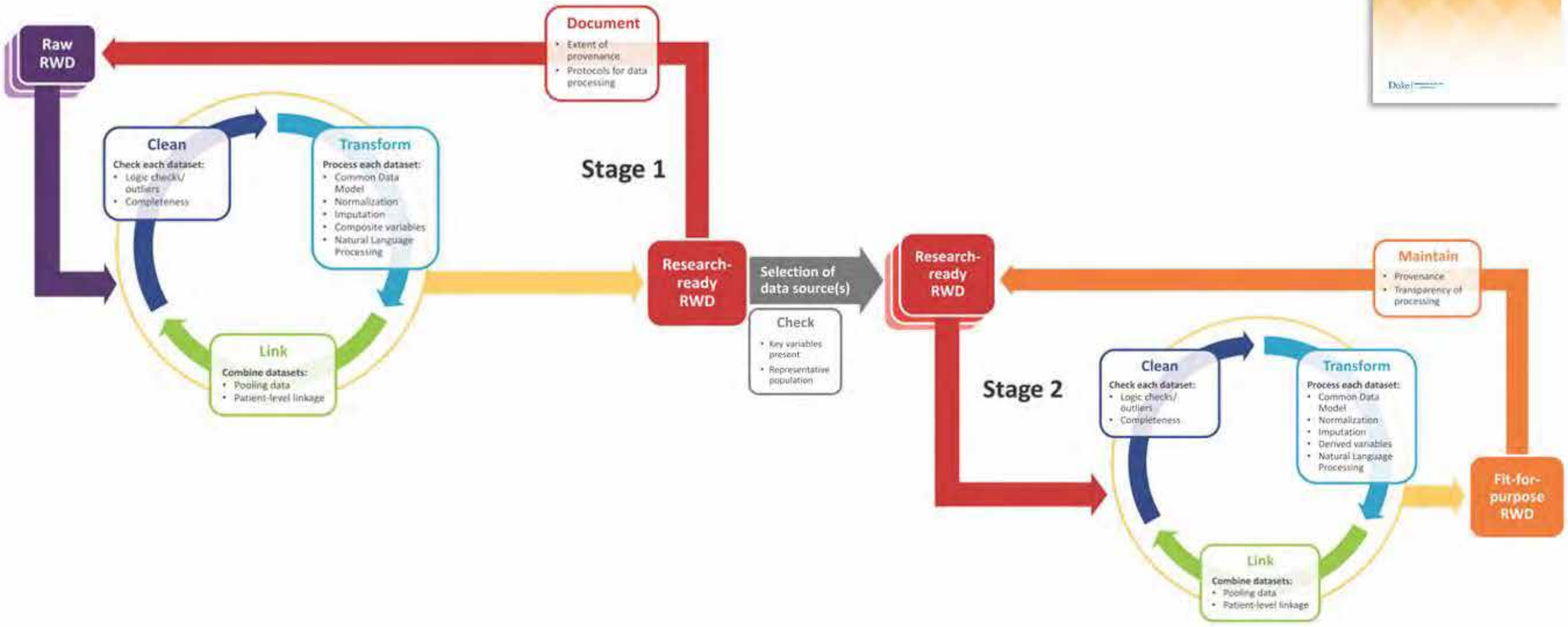
Sentinel Quality Review and Characterization Learnings

- NDC codes in the Procedure Table
  - SOC: There is a significant change in the number of records where the PX variable has values with special characters other than a decimal point across ETLs
  - Response: due to the addition of the claimline NDC to the PX variable; These NDC values are not adjudicated like pharmacy claims so often contain dash elements of the NDC code.
- ICD9 diagnosis/procedure codes post October 2015
- Claims before birthdate/Claims after deathdate
- Multiple patids for same members(kids 0-2) found while mom-baby linkage

Query Specific Quality Review and Characterization

- Cohort Identification and Descriptive Analysis Module (CIDA) identifies and extracts cohorts of interest from the Sentinel Distributed Database based on user-defined options
  - Exposures, outcomes, continuous enrollment requirements, incidence criteria, inclusion/exclusion criteria, relevant age groups, demographics
- Data Partners review output from CIDA requests to check for population plausibility, program errors, code list omissions
  - Data ambiguity from the time around birth when the infant does not have his or her own member number, so claims are submitted under the mother's ID
  - Review of code lists with internal NDC resources within specific queries or review of claim lines for code modifiers (biologics and biosimilars)

Curation is Often Complex and Hard to Explain



Duke | MARGOLIS CENTER for Health Policy

Session I: Establishing a High-Quality RWD Ecosystem

Break

Session II: Curating and Assessing Fit-for-Use RWD Derived from Electronic Health Records

Curation of EHR data

Keith Marsolo, PhD
Associate Professor
Department of Population Health Sciences
Duke Clinical Research Institute
Duke University School of Medicine

pcornet®

The National Patient-Centered Clinical Research Network

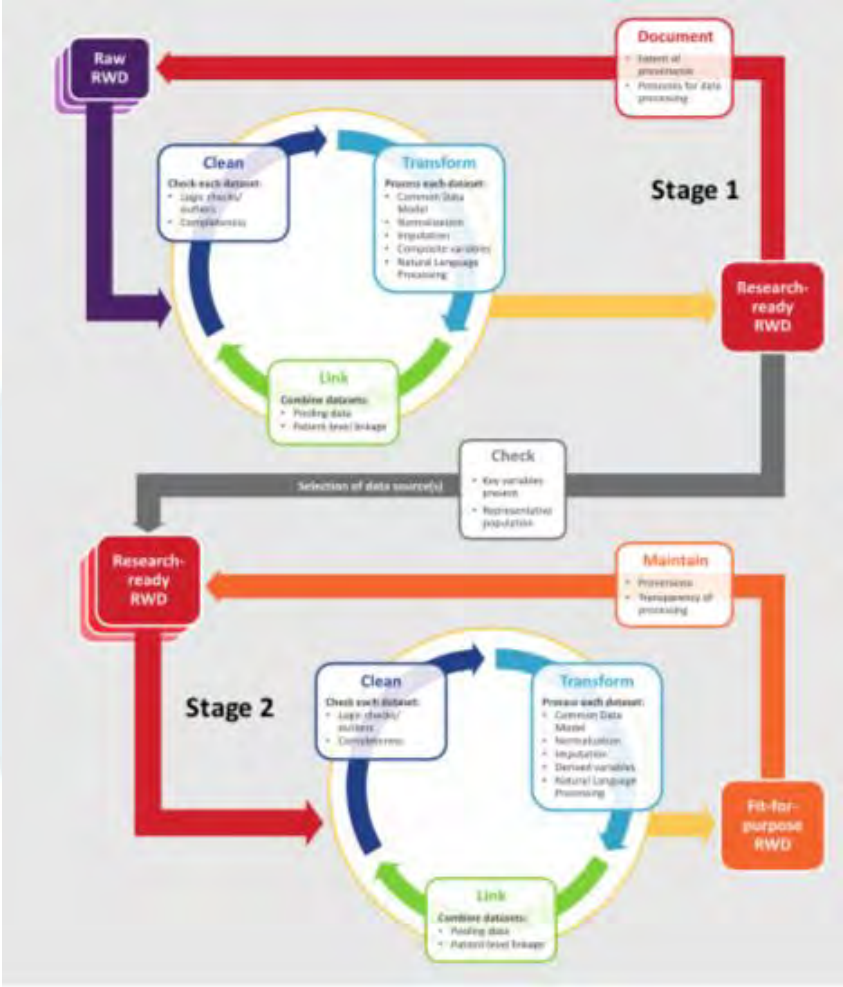
Disclosures

- Consulting support from Novartis & IBM
- Co-inventor – Hive Networks, Inc.

pcornet

Moving from raw data to fit-for-purpose – PCORnet®

- PCORnet follows a two-stage process to assess suitability
  - **Foundational** curation – establish a baseline level of data quality (**"minimum necessary"**)
  - **Study-specific** – ensure data are fit-for-purpose for a given study or analysis
- Foundational data curation is not static – view as a **continuous learning cycle**
  - Continuous assessment of performance
  - Close gap between foundational and study-specific – add new data checks based on study findings

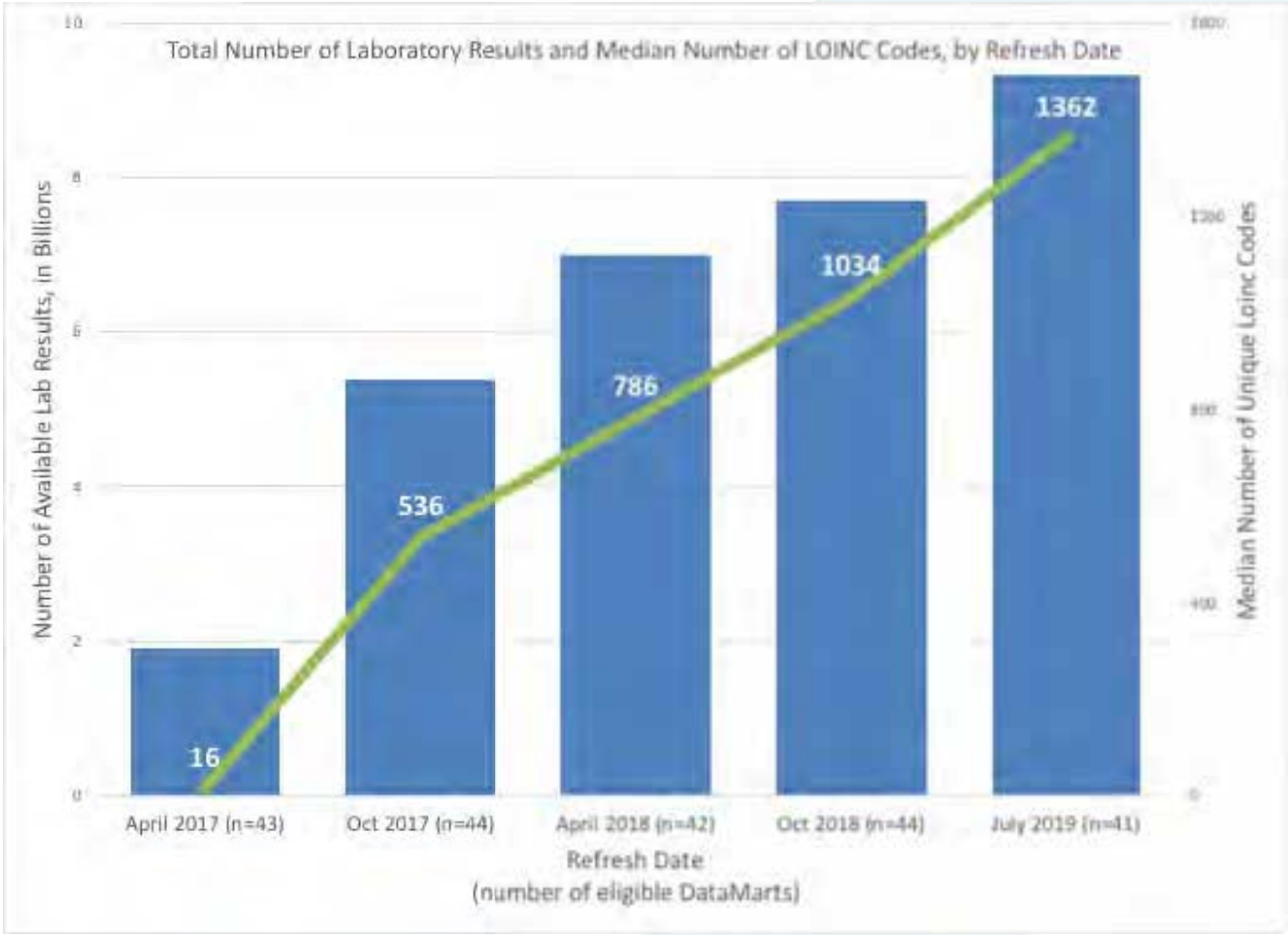


pcornet

Why foundational curation?

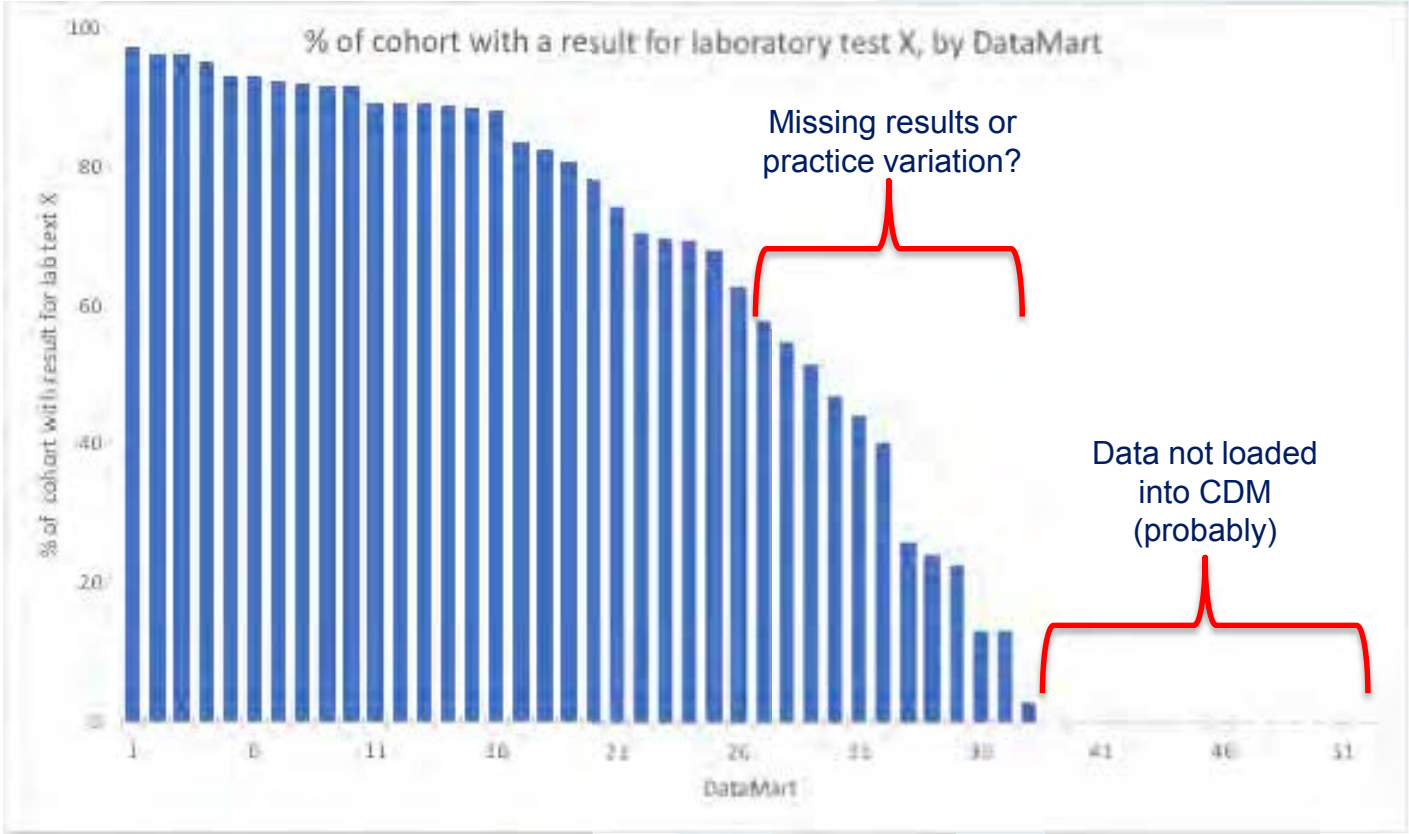
- Many EHR domains are being harmonized / standardized for the first time
- Given volume of data, can be overwhelming to both harmonize and assess fitness for specific study questions at the same time

Selected lab-related data checks (failure criteria)
Less than 80% of lab results mapped to LOINC
Less than 80% of quantitative lab results specify the normal range
Less than 80% of quantitative lab results mapped to LOINC specify specimen source & result unit
More than 5% of lab results have inappropriate specimen source [for selected tests]
Median lab result values for selected tests are statistical or clinical outliers



Study specific curation

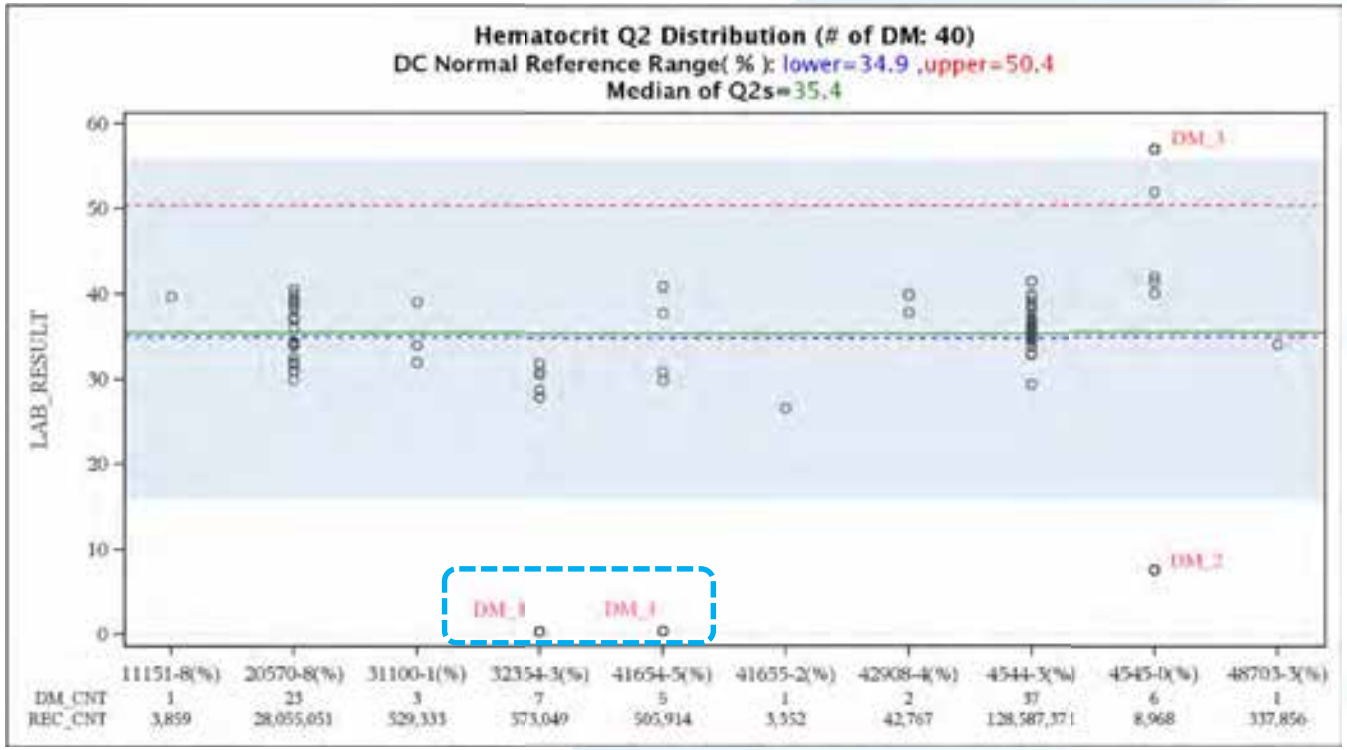
- Identify potential quality concerns for key variables within a given study populations
- Need to determine whether issues are related to the data or reflect normal practice variation



Minimum necessary data checks

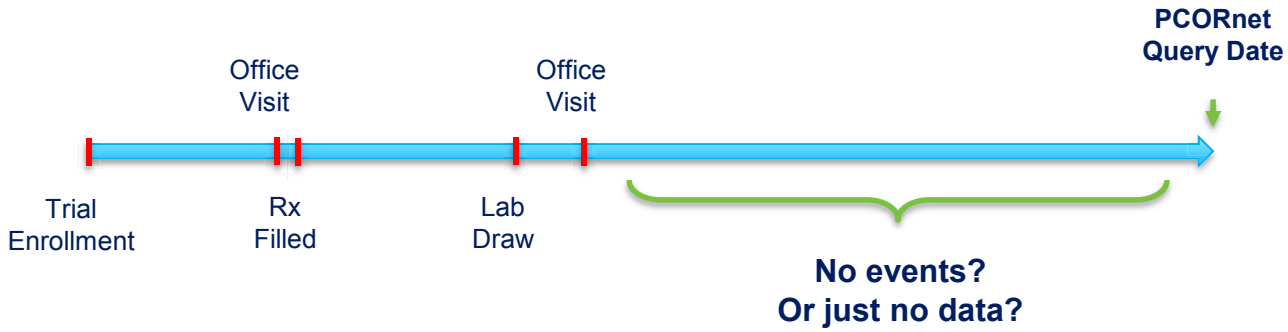
- Need to align checks with purpose
- Will data be confirmatory, or serve as stand-alone outcome / endpoint
- If minimum threshold cannot be met, can dataset be used for something else?

Selected lab-related data checks (failure criteria)	% of DataMarts passing (most recent refresh; n=41)
Less than 80% of lab results mapped to LOINC	85%
Less than 80% of quantitative lab results specify the normal range	42%
Less than 80% of quantitative lab results mapped to LOINC specify specimen source & result unit	37%



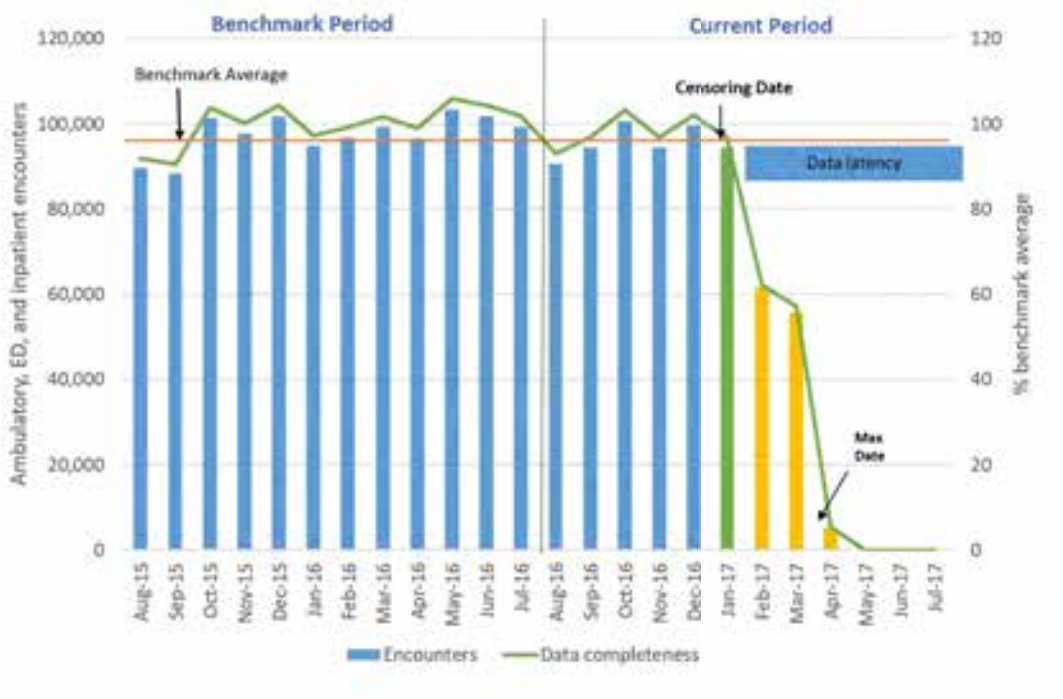
Curation as a learning process – data latency

- Latency / completeness of data

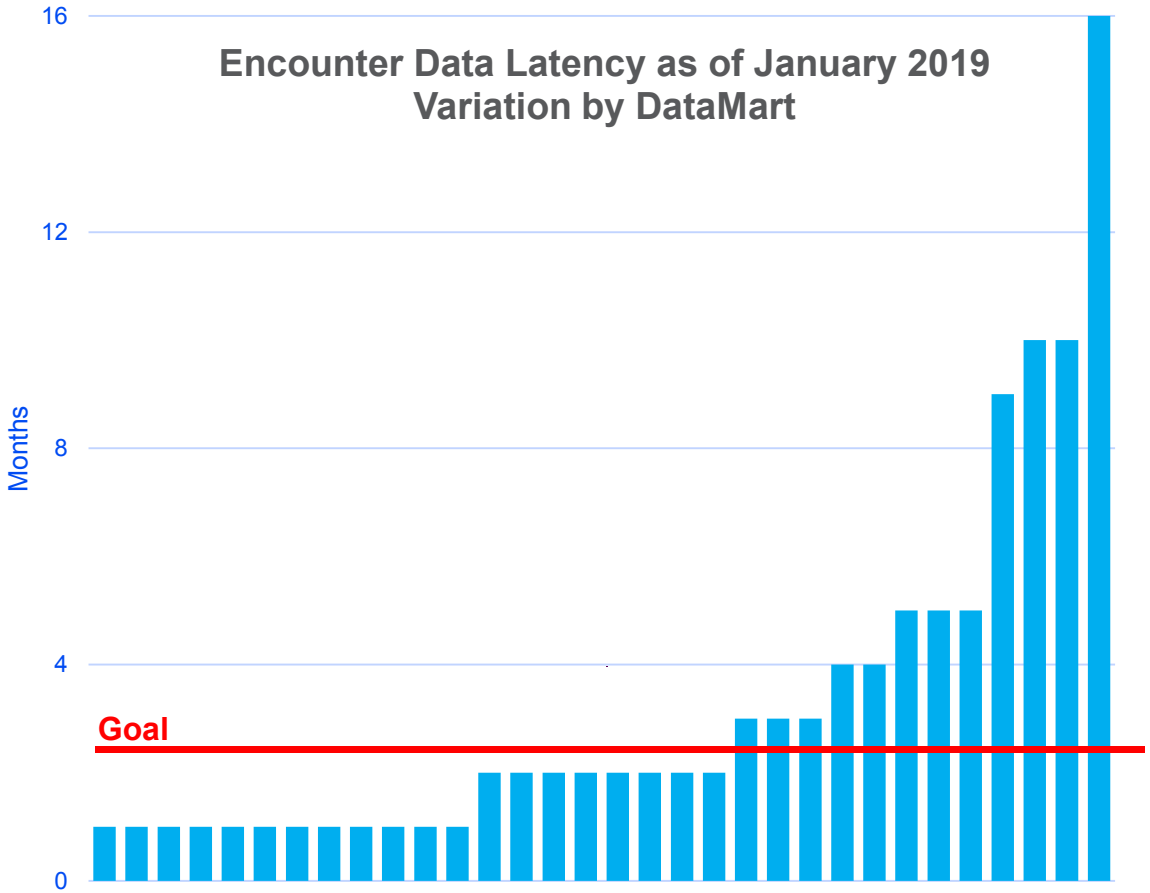


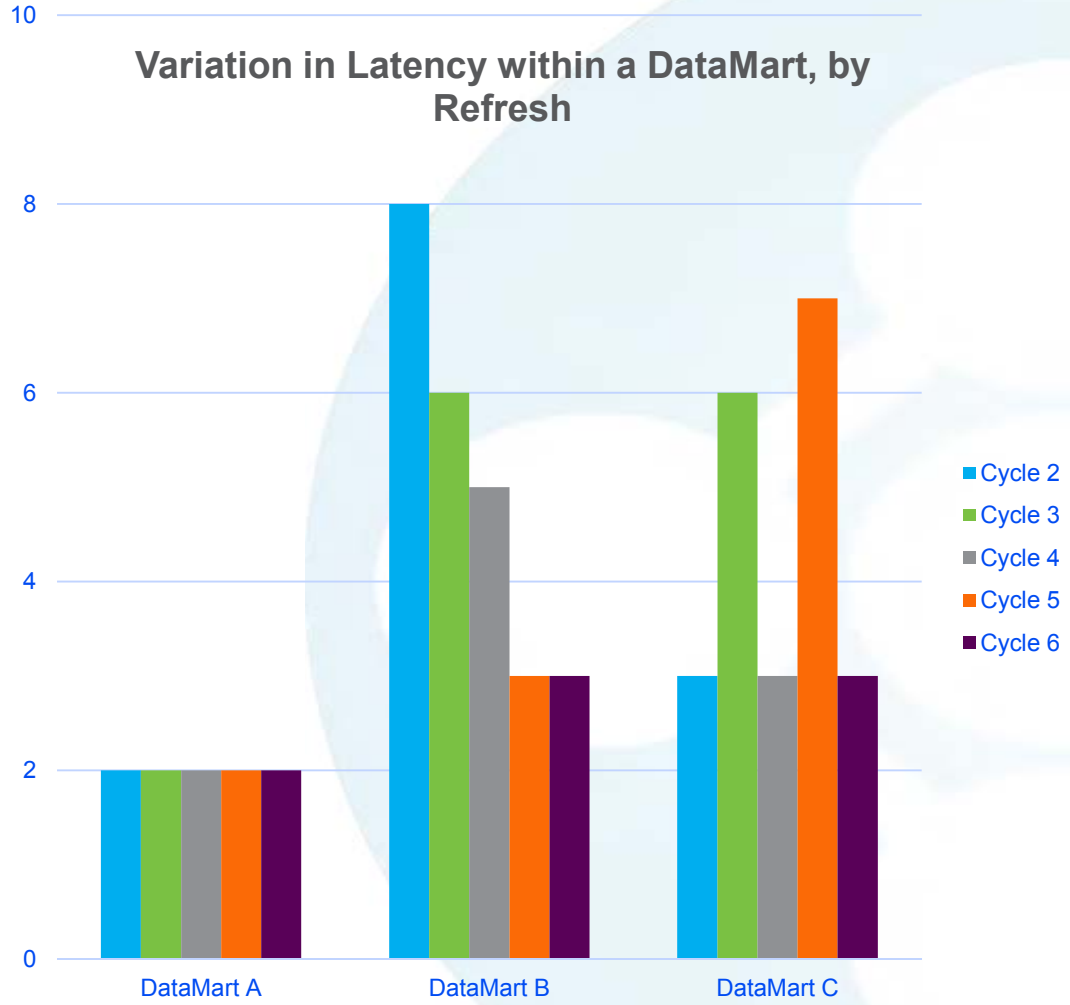
- Questions:
  - *"How complete & up-to-date are the data?" (DSMB)*
  - *"What's the data censoring date for participants?" (Statistician)*

- Developed latency calculation & incorporated into data curation



Curation as a *continuous* learning process





Eligible DataMarts: PCORnet 2.0 DataMarts which include inpatient, ambulatory, and/or Emergency Department encounters and do not use date obfuscation

Summary

- Data curation should be viewed as a process for continuous quality improvement
- May not end up with a single set of “minimum necessary” checks – consider tiered approach
- As best practices are developed, need a better way to share methods, results, etc.
- Have spent years understanding the pitfalls of working with administrative claims – will take time to develop that knowledge around EHR data

pcornet

Session II: Curating and Assessing Fit-for-Use RWD Derived from Electronic Health Records

Lunch

Session III: Leveraging Digital Technology for Patient-Generated Health Data

evidation

Understanding PGHD Data Quality in the Real World

Ernesto Ramirez, PhD
Senior Data Scientist
Evidation Health, Inc.

eramirez@evidation.com
@eramirez

Developing Real-World Data and Evidence to Support Regulatory Decision-Making | Duke-Margolis Center for Health Policy

OCTOBER 3, 2019

Outline

Background: Person-Generated Health Data (PGHD)

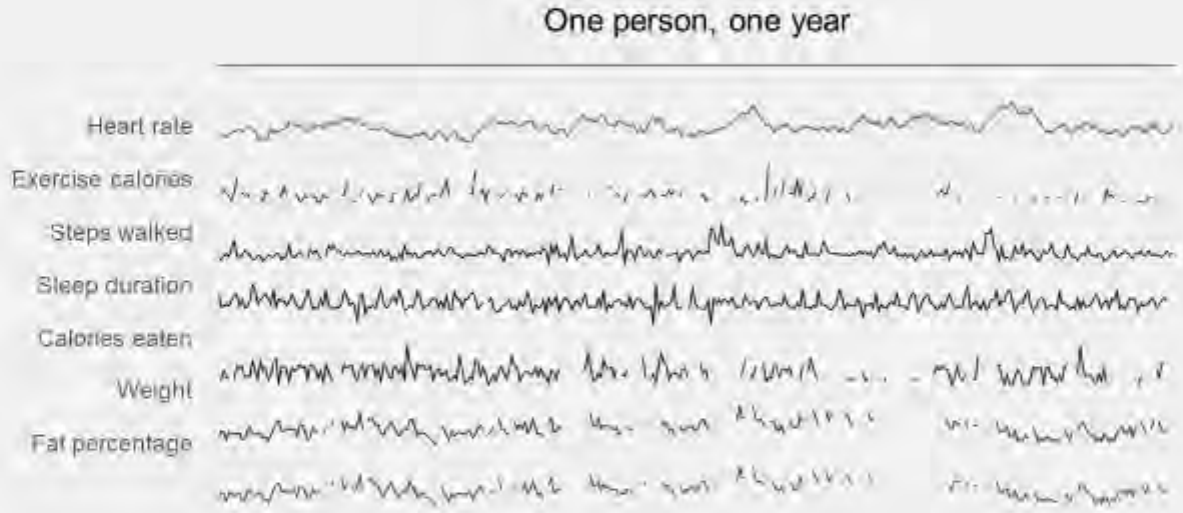
Case Study: Developing Measures of Cognitive Impairment in the Real World from Consumer-Grade Multimodal Sensor Streams

Data Quality: Five considerations for PGHD

Person-Generated Health Data (PGHD) enables continuous monitoring of health outcomes at the individual level so we can better understand and measure a person's experience.

OBJECTIVE EVERYDAY DATA

Collected via sensors and apps



PHENOTYPIC LABELS

Collected via questionnaires

Some values can be verified via traditional data sources (e.g., claims, labs, EHR)

Age | Gender | Ethnicity / race | Medical diagnoses | Prescription drugs | Smoking

Education | Household | Zip code | Patient-reported outcomes | Fast food consumption

Employment status | Employer | Supplements | Quality of life | Alcohol use

Insurance carrier | Height / weight | Major medical events | Sleep quality

EXAMPLE DIGITAL ASSAYS

Direct assessment
Functional mobility
Sleep reliability
Weight range

Behavioral inference
Routine/consistency
Digital utilization
Responsiveness

Clinical inference
Exacerbation events
Treatment utilization
Disease progression

PGHD allows for measuring novel outcomes for chronic conditions at the population level.

	People w/ T2DM	Matched controls
Number of activity trackers	4,459	10,321
% days with tracked steps*	78.7%	80.7%
Mean nightly sleep duration (hours)*	6.48	6.69
Sleep regularity index (SRI)*	0.72	0.77
Resting heart rate (BPM)*	71.2	66.0

*p < 0.05

SOURCE: USING CLAIMS AND WEARABLE DEVICES DATA TO QUANTIFY INFLUENZA OUTCOMES AMONG TYPE 2 DIABETES PATIENTS - A POPULATION STUDY. SAMSON ET AL., AMERICAN DIABETES ASSOCIATION, JULY 2018

	People w/ MS	Matched controls
Number of activity trackers	498	1,400
% days with tracked steps**	73%	77%
Mean daily step count**	6,379	7,188
Mean nightly sleep duration (hours)	6.3	6.5
Max time to fall asleep (minutes)**	18.58	13.91

*p < 0.05 **p < 0.001, FDR-ADJUSTED

SOURCE: REAL-WORLD USE OF WEARABLE DEVICES IN A LARGE MULTIPLE SCLEROSIS COHORT. FOSCHINI ET AL., AMERICAN ACADEMY OF NEUROLOGY, APRIL 2018

Outline

Background: Person-Generated Health Data (PGHD)

Case Study: Developing Measures of Cognitive Impairment in the Real World from Consumer-Grade Multimodal Sensor Streams

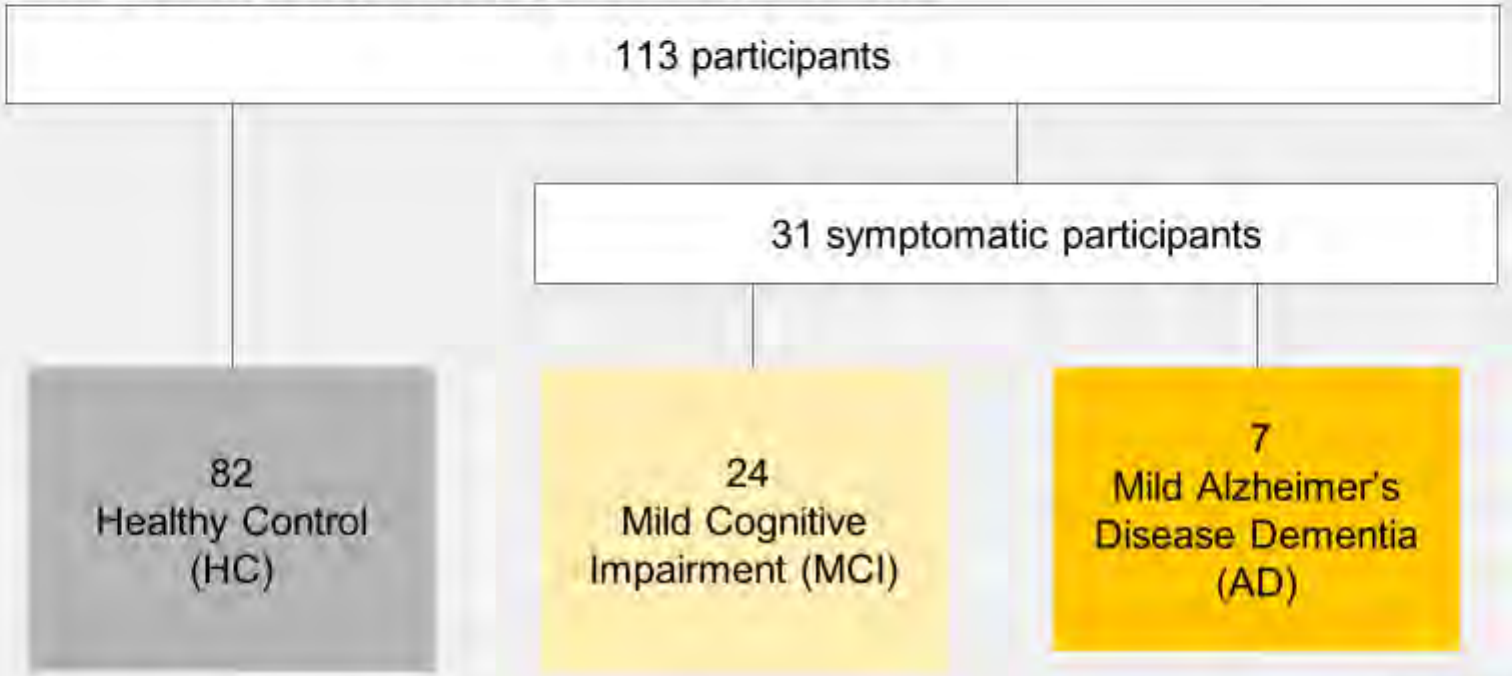
Data Quality: Five considerations for PGHD

Evidation, along with collaborators at Eli Lilly and Apple, recently completed a study using PGHD in participants with cognitive decline.

Objectives

1. Assess the feasibility of collecting and processing data from multiple smart devices of older adults with and without cognitive impairment in their daily lives.

2. Test whether data from these devices can differentiate between healthy controls and participants with cognitive impairment.



DOI.ORG/10.1145/3292500.3330690

Participants were given an iPhone, Apple Watch, and Beddit sleep monitor to use as their primary devices over the course of the 12 week study, as well as an iPad to complete at-home cognitive tests.

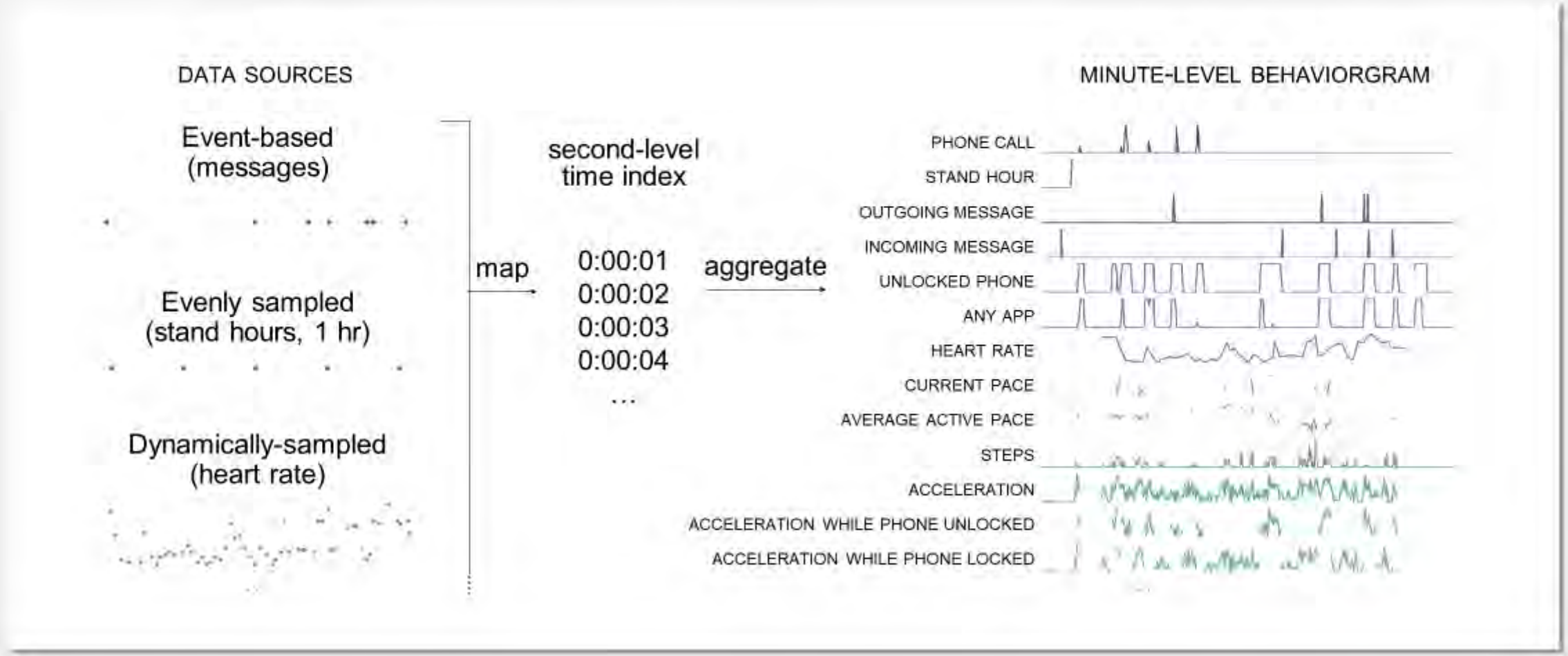
ACCELERATION
STEPS
CALLS
MESSAGES
APP USAGE
DISTANCE
DAILY SURVEYS

ACCELERATION
STEPS
STANDING
HEART RATE
EXERCISE

SLEEP

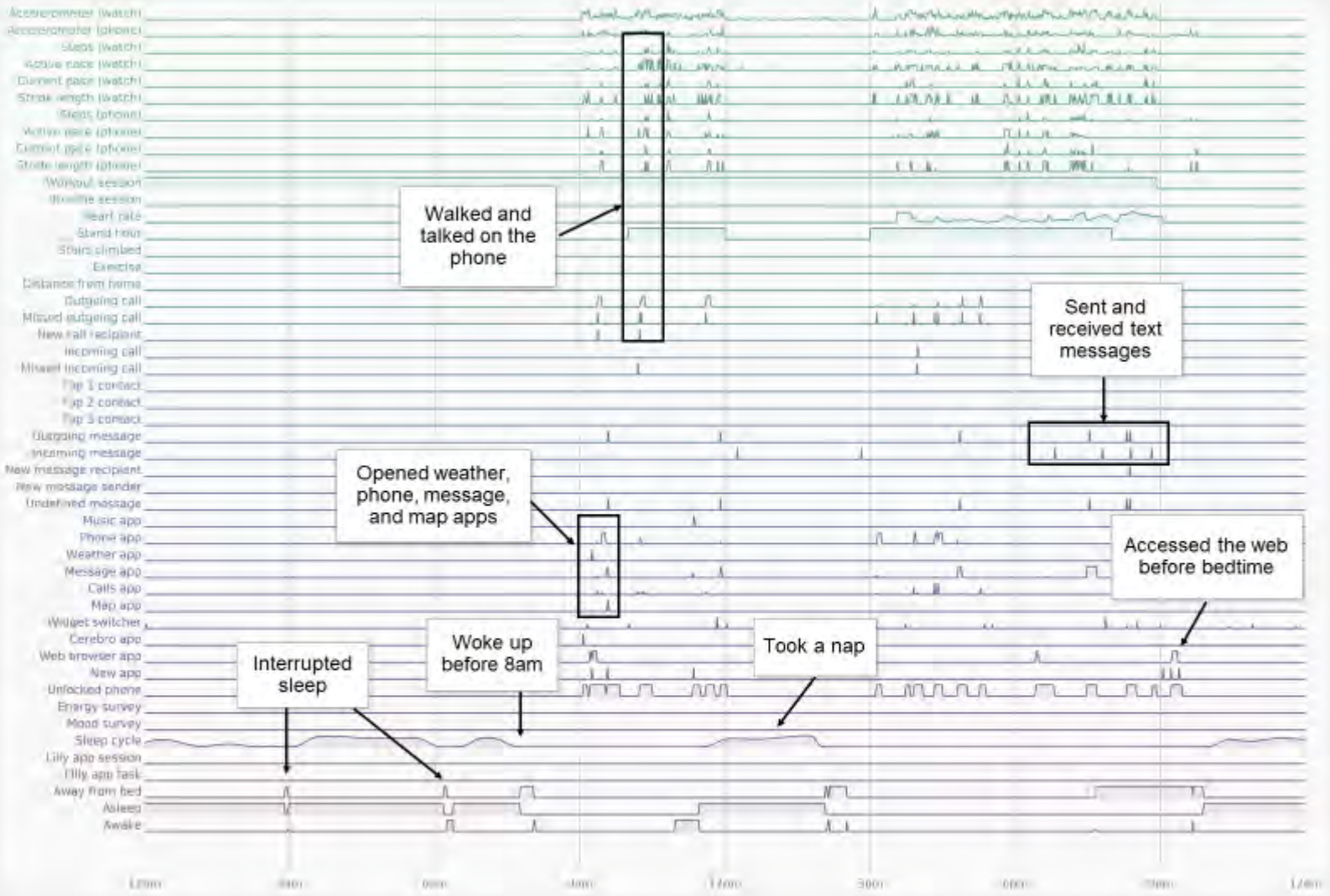
PSYCHOMOTOR SKILLS
COGNITIVE PERFORMANCE

We processed, aligned, and combined data from all the different data sources to create a single behaviorgram for each participant.



The behaviorgram offers a rich representation of an individual's behavior.

It also serves as tool for data exploration, hypothesis generation, and most importantly, a way inspect the quality of the data.



Outline

Background: Person-Generated Health Data (PGHD)

Case Study: Developing Measures of Cognitive Impairment in the Real World from Consumer-Grade Multimodal Sensor Streams

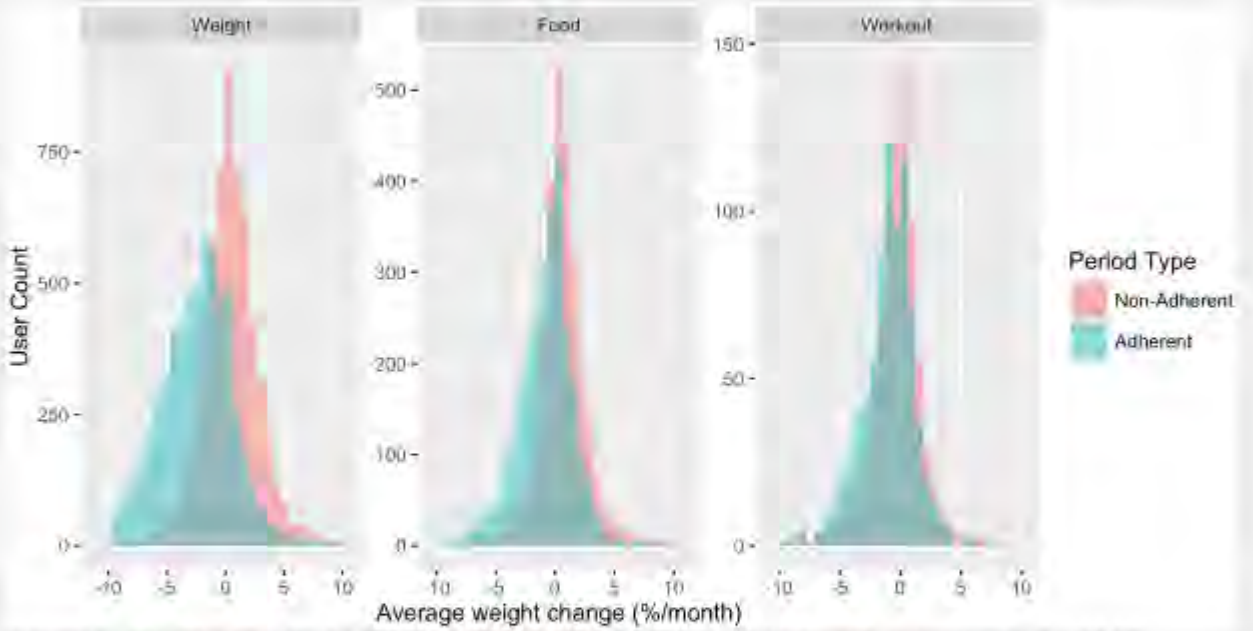
Data Quality: Five considerations for PGHD

One: Understand and characterize your data, then determine reasons for observed issues with collected data.

1. Identify and characterize issues in data by using data coverage tools and reports.
   - Use an aligned and standardized resolution to produce visualizations for quick exploration.
2. Determine if issues are systematic (due to device sensors, data collection or ingestion) or behavioral.
   - Understanding system architecture and data flows is crucial for developing data quality checks.
   - Real-world data means encountering real-world problems.

Two: Develop and implement replicable methods for dealing with issues.

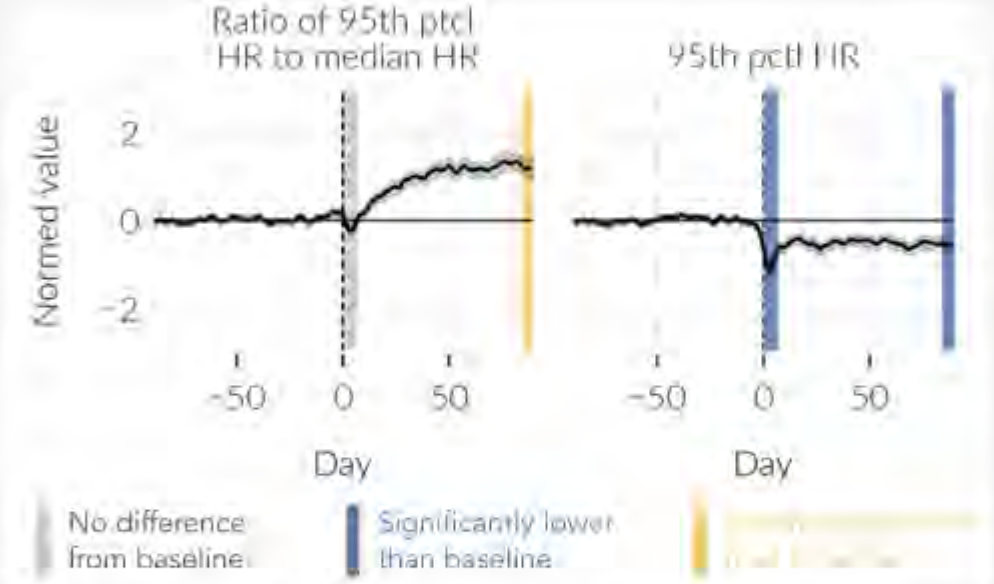
1. Systematic issues can be addressed using appropriate imputation techniques.
   - Sparse sampling? Impute to fill the gaps
   - Use cross-channel information to determine the correct strategy.
2. Issues due to behavioral factors need further exploration for possible inclusion.
   - Missingness can be an informative feature in many situations.



SOURCE: ADHERENT USE OF DIGITAL HEALTH TRACKERS IS ASSOCIATED WITH WEIGHT LOSS. POURZANJANI ET AL., PLOS ONE 2016

Three: Apply appropriate analysis methods that accurately characterizes the outcomes of interest

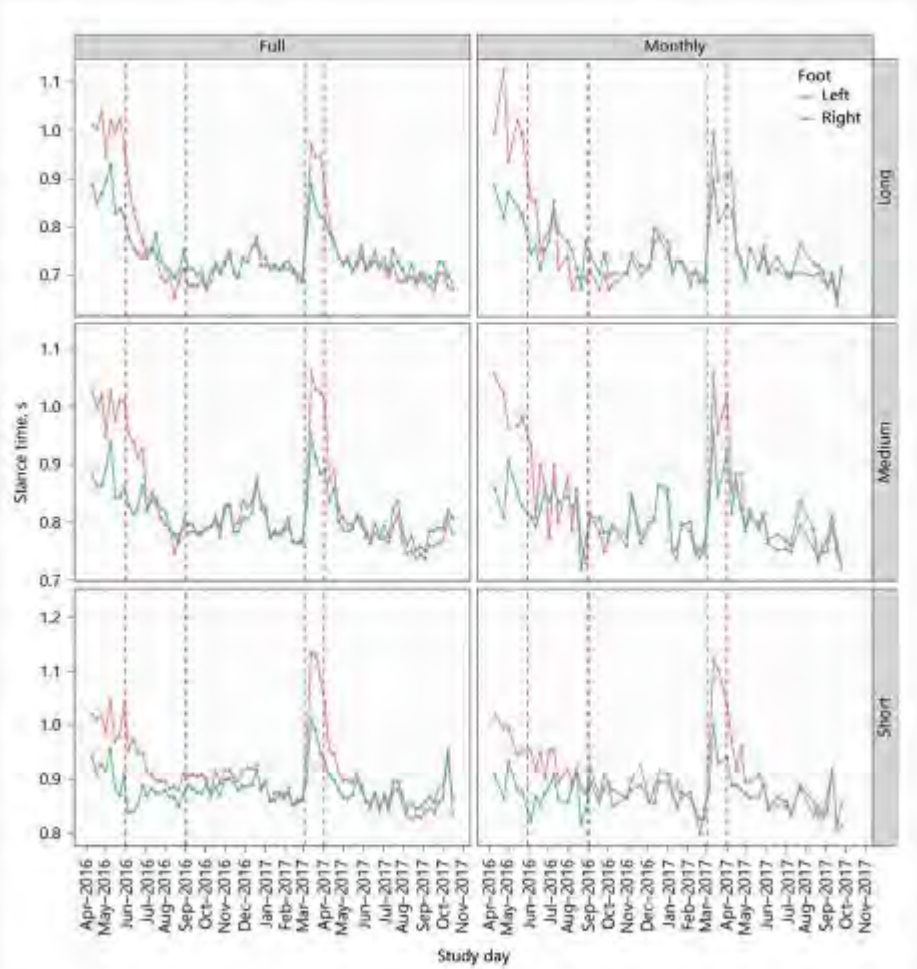
1. PGHD will typically include outliers that may reflect true observations.
   - Real-world data capture needs to account for data collection issues and behavioral artifacts.
   - Important to thoroughly investigate outliers for plausibility.
2. Use statistical aggregations that are robust to outliers. For example:
   - ~~Mean~~ Median
   - ~~Max~~ 95th Percentile
   - ~~Standard Deviation~~ Interquartile Range



SOURCE: CONTINUOUS DIGITAL ASSESSMENT FOR WEIGHT LOSS SURGERY PATIENTS. RAMIREZ ET AL., IN REVIEW.

Four: Test endpoint(s) for sensitivity to potential issues with data quality.

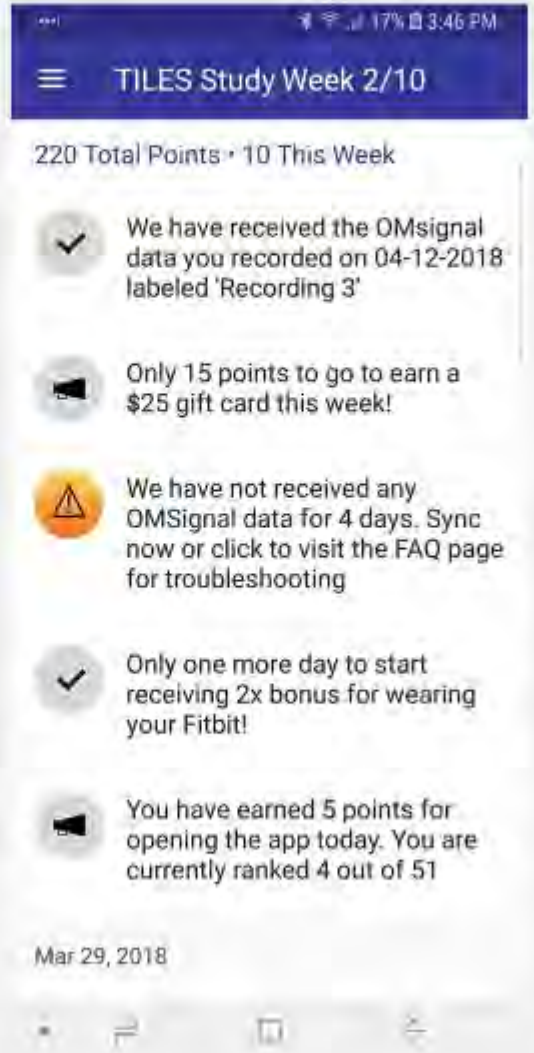
1. Are the endpoints robust to varying amounts of available data / compliance?
   - Resample data to simulate changes in data availability and evaluate for minimum required data.
   - What is the minimum amount of data need to generate sound inferences?



ADAPTED FROM CONTINUOUS MONITORING OF PATIENT MOBILITY FOR 18 MONTHS USING INERTIAL SENSORS FOLLOWING TRAUMATIC KNEE INJURY: A CASE STUDY. MUELLER ET AL., DIGITAL BIOMARKERS 2018

Five: Use features of continuous data streams to evaluate and improve data quality in real-time.

1. Capitalize on data availability to build real-time quality checks.
   - Completeness and conformance checks do not need to wait until data collection is finalized.
   - Data flows can be checked against minimum standards related to compliance and plausibility.
2. Use data as a feedback mechanism to involve participants in data quality process.
   - Real-time feedback can improve participant engagement.



evidation

Ernesto Ramirez, PhD
Senior Data Scientist


eramirez@evidation.com


@eramirez

Session III: Leveraging Digital Technology for Patient-Generated Health Data

Integrating Multi-Dimensional Real World Data to Accelerate Research and Enhance Patient Centricity

Angela Dobes, MPH
Senior Director, IBD Plexus

IBD PLEXUS

CROHN'S & COLITIS FOUNDATION

IBD Plexus is designed to support

Discovery

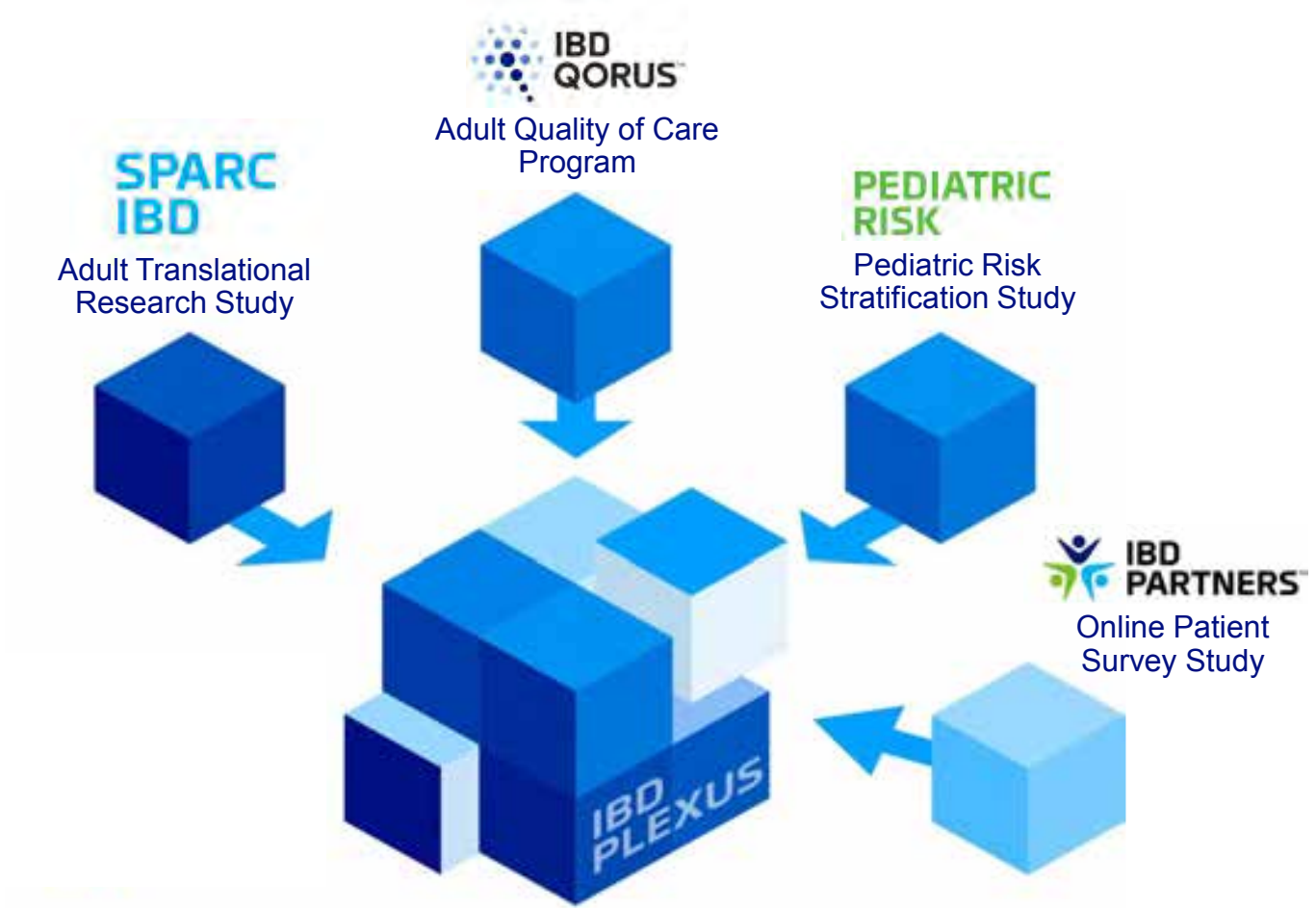
Clinical Development

Post Approval

IBD PLEXUS

CROHN'S & COLITIS FOUNDATION

Diverse research cohorts for cutting edge research



Real-world data integrated & linked within & across cohorts

Patient surveys

Electronic case report forms

Labs

Molecular data

Medical record

PRIMARY RWD

- Patient reported data
- Clinician reported data

IBD SmartForm

SECONDARY RWD

IBD PLEXUS

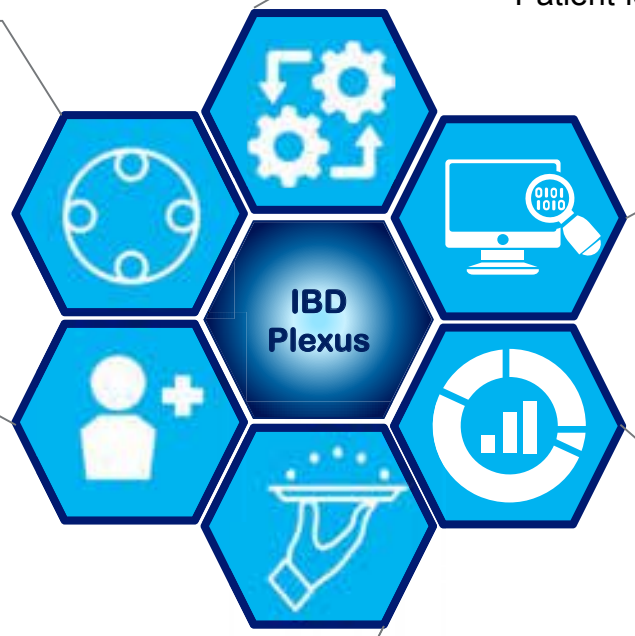
Information Management Lifecycle

Integration & linkage

- Data integration engine & processing tools
- Master patient index engine
- Patient-level linkage

Standardization & normalization

- Data collection standards & protocols
- Common data models
- Data harmonization tools



Quality control

- Built-in data profiling & error reporting mechanisms
- Balance between system & manual data checks
- Processing history controls

Registration & authentication

- Master consent / HIPAA authorization
- Multi-study registration functionality
- Patient re-contact capabilities

Prep-to-research tools

- Data querying capabilities
- Data visualization and insight tools

Data delivery

- Automated data provisioning process
- Raw & research-ready datasets
- Data dictionary
- White glove service

Achieving Research & Development Efficiencies with RWD

4 Research study cohorts

Over 70 participating sites

8 Pharmaceutical companies

3 Ancillary study awards *(CDC, NIH, PCORI)*

- Hypothesis generation
- Drug development tools
- Study feasibility & recruitment
- Identification of characteristics for enrichment or stratification

IBD PLEXUS

CROHN'S & COLITIS FOUNDATION

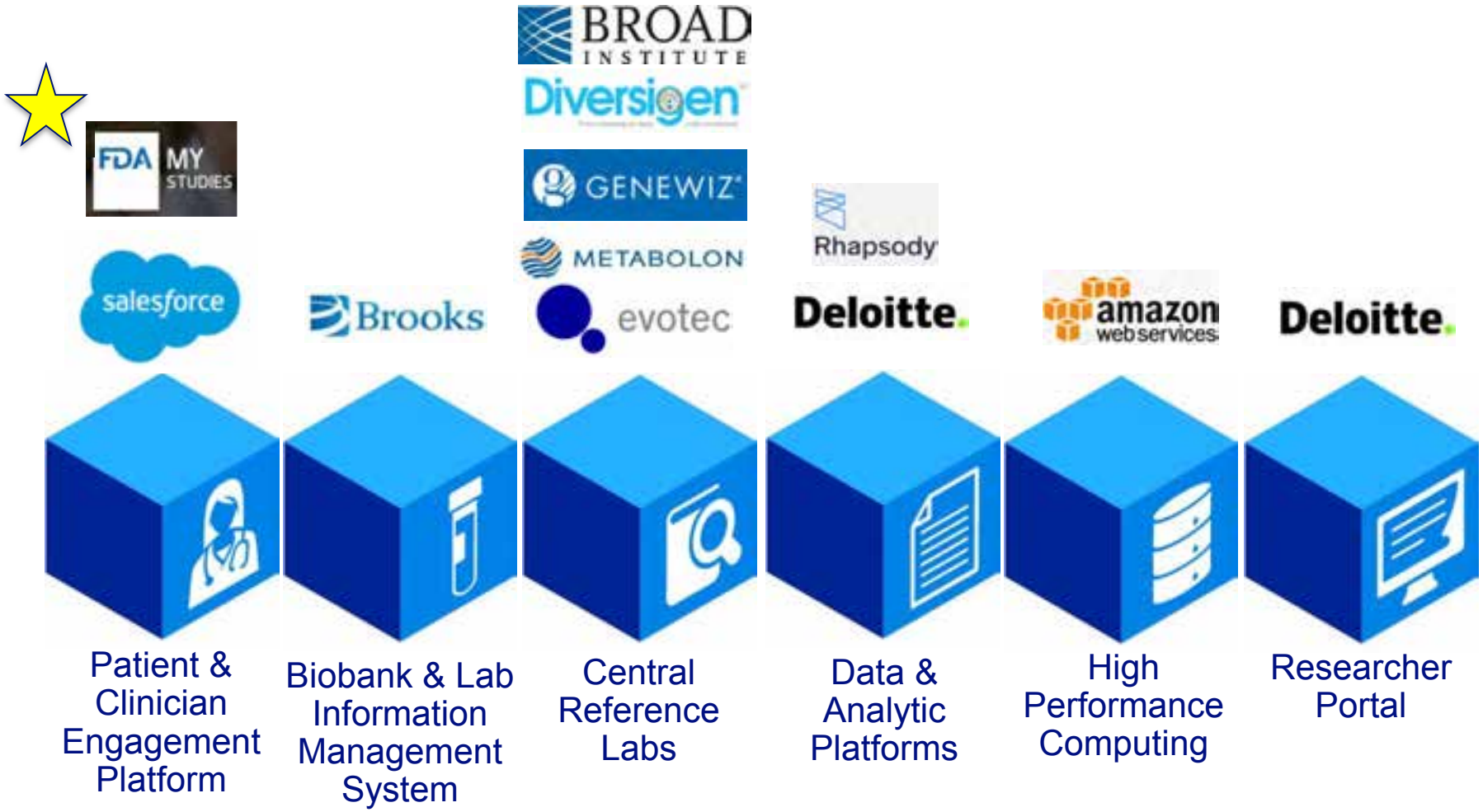
Mindful of the patient journey, we embrace a patient-centric approach to all decision-making and mission delivery.

FDA Real-World Evidence Program Demonstration Project

IBD PLEXUS

CROHN'S & COLITIS FOUNDATION

Powering IBD Plexus



IBD PLEXUS

CROHN'S & COLITIS FOUNDATION

Demonstration Project Goals

- Explore the use of a digital mobile app to fill data gaps
- Capture the patient experience beyond the clinical delivery system
- Establish a more comprehensive picture how medical products function beyond the controlled restrictions of traditional randomized clinical trials
- Help assess the use of patient-generated health data to support RWE

Establishing a High-Quality PGHD Ecosystem:

Patient reported data considerations

Completeness

- Disease activity
- UX barriers
- Life events

Conformance

- Standards
  - Instruments
  - Mechanisms
- Collection windows

Credibility

- Validation
- Recall bias
- Business rules

Error profiling reports & quality checks

IBD PLEXUS

CROHN'S & COLITIS FOUNDATION

Session III: Leveraging Digital Technology for Patient-Generated Health Data

Session IV: Methodological and Analytical Considerations for Observational Studies

Session IV: Methodological and Analytical Considerations for Observational Studies

Where Have We Come From – Where Are We Now – Where Are We Going?

Til Stürmer, MD, MPH, PhD

October 3rd, 2019

Developing Real-World Data and Evidence to Support Regulatory Decision-Making

National Press Club • Washington, DC

ispe International Society for Pharmacoepidemiology

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

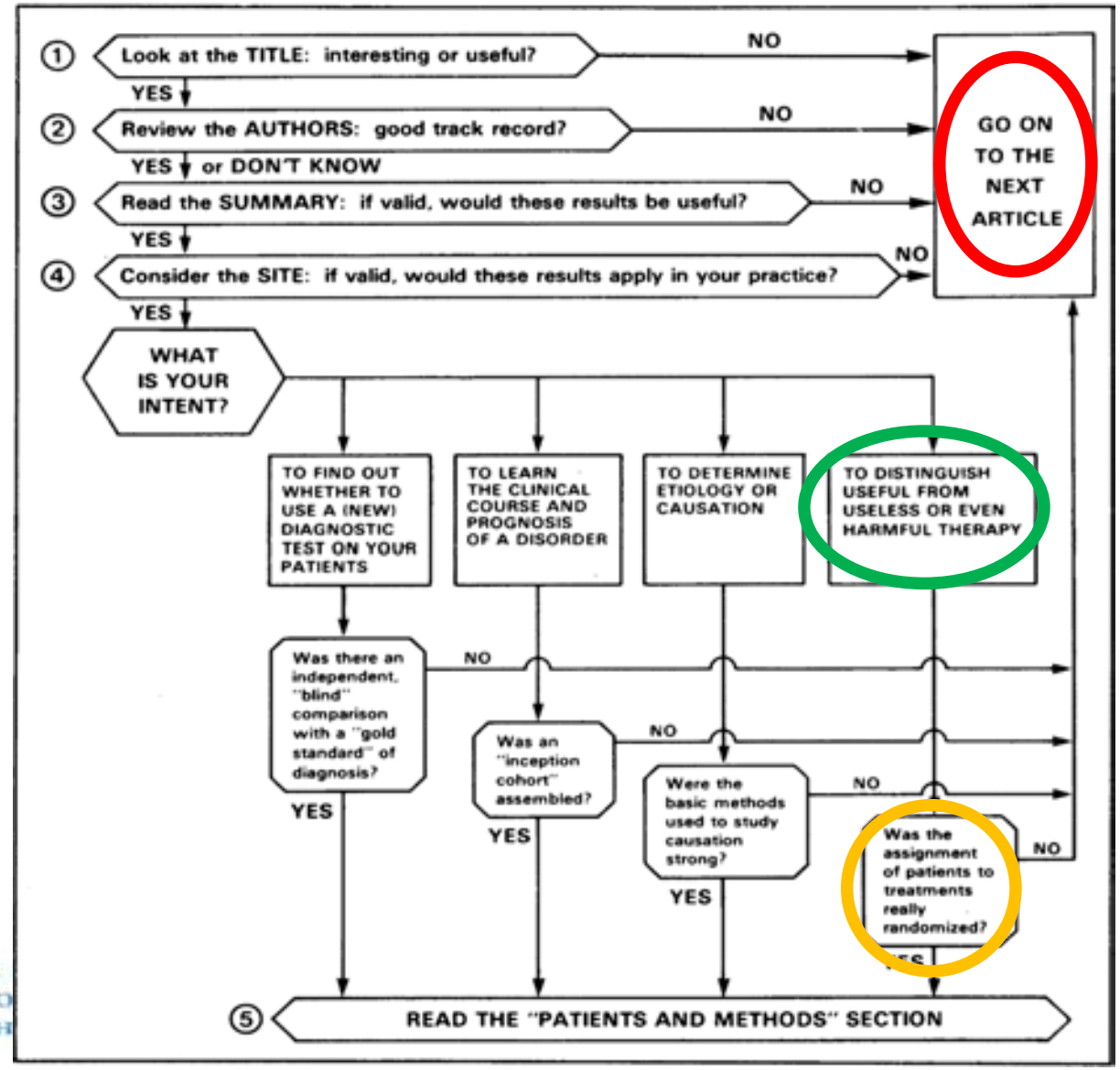
Disclosures

The following personal or financial relationships relevant to this presentation existed during the past 12 months:

- I receive investigator-initiated research funding and support as Principal Investigator (R01 AG056479) from the National Institute on Aging (NIA), and as Co-Investigator (R01 HL118255, R01MD011680), National Institutes of Health (NIH)
- I receive salary support as Director of Comparative Effectiveness Research (CER), NC TraCS Institute, UNC Clinical and Translational Science Award (UL1TR002489), from the Center for Pharmacoepidemiology (current members: GlaxoSmithKline, UCB BioSciences, Merck, Takeda), from pharmaceutical companies (Novo Nordisk), and from a generous contribution from Dr. Nancy A. Dreyer to the Department of Epidemiology, University of North Carolina at Chapel Hill.
- I do not accept personal compensation of any kind from any pharmaceutical company
- I own stock in Novartis, Roche, BASF, AstraZeneca, and Novo Nordisk.

Where Have We Come From?

Intractable Confounding



Sackett DL. How to read clinical journals: I. Why to read them and how to start reading them critically. CMAJ 1981

Miettinen Stat Med 1983: ”control of the indication … commonly infeasible”

Yusuf, Collins, Peto. Stat Med 1984: “little real value“

Confounding by Indication

- Good prescribing leads to confounding of drug effects on intended outcomes
- More severe disease more likely to
  - Be treated (with higher doses)
  - Have higher risk of adverse outcomes
- Assessment of severity of disease
  - Often difficult
  - Intractable for intended effects (Miettinen 1983; Yusuf, Collins, & Peto 1984)
- Drug looks BAD compared with NON-USERS!

Confounding by Frailty

- Individuals close to death are
  - Less likely to receive preventive treatments
    - E.g., statins, flu vaccination
  - More likely switched to palliative treatments
    - E.g., opiates instead of NSAIDs
  - More likely to receive certain classes of drugs
    - E.g., loop diuretics vs. other diuretics
- Paradoxical drug mortality associations
- Drug looks GOOD compared with NON-USERS!

Ignoring Adherence and Time on Treatment

Epidemiology

Statin Adherence and Risk of Accidents

A Cautionary Tale

Colin R. Dormuth, ScD; Amanda R. Patrick, SM; William H. Shrank, MD;
James M. Wright, MD, PhD; Robert J. Glynn, PhD, ScD;
Jenny Sutherland, BSc; M. Alan Brookhart, PhD

Format: Abstract ▾ Send to ▾

Am J Epidemiol. 2019 Mar 27. pii: kwz066. doi: 10.1093/aje/kwz066. [Epub ahead of print]

Keeping the Demons At Bay When Handling Time Varying Exposures: Beyond Avoiding Immortal Person Time.

Edwards JK[1], Htoo PT[1], Stürmer T[1].

Author information

1 Department of Epidemiology, University of North Carolina at Chapel Hill.

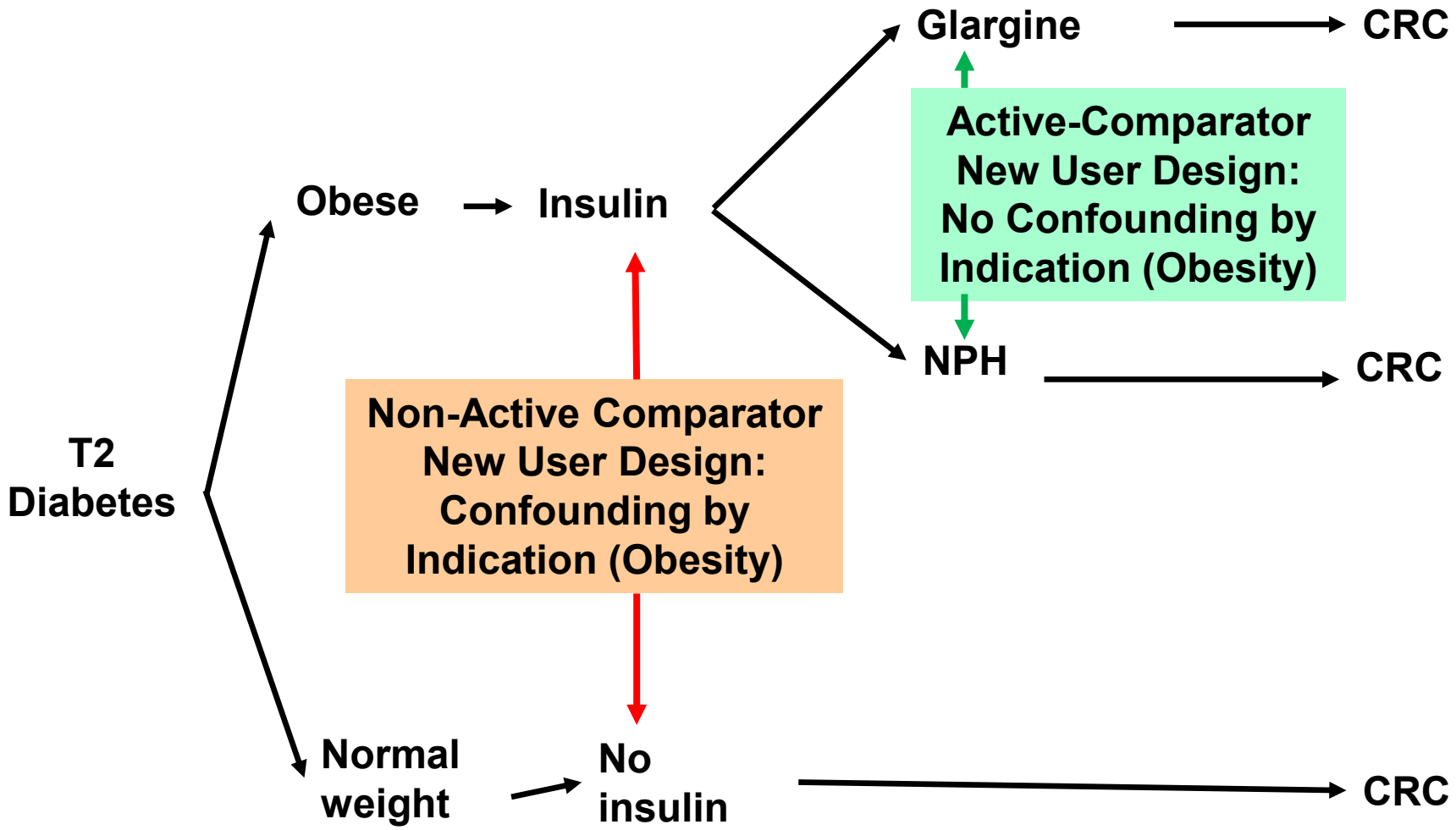
THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

Conclusions: Where Have We Come From?

- Comparison of prevalent drug users to non-users standard study design until ~15 years ago
- Suffers from all these biases: Often invalid
- Statements about validity of nonexperimental research based on such comparisons/designs
- There may be few exceptions where biases work in our “favor” (e.g., long latent period)

Where Are We Now?

Active Comparator, New User Design



Confounding Control by Design: BMI

Table 4—*Effect of BMI on channeling between initiating glargine versus initiating NPH: external validation studies*

	Glargine	NPH
MGH		
n	574	412
BMI (kg/m^2), mean ± SD*	32.7 ± 7.53	32.4 ± 8.43
BMI (kg/m^2), *n* (%)		
<19	4 (0.7)	8 (1.9)
19 to <25	77 (13.4)	67 (16.3)
25 to <30	150 (26.1)	105 (25.5)
30 to <35	146 (25.4)	104 (25.2)
35 to <40	114 (19.9)	64 (15.5)
40 to <45	45 (7.8)	36 (8.7)
≥45	38 (6.6)	28 (6.8)

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

Stürmer T, Marquis MA, Zhou H, Meigs JB, Lim S, Blonde L, MacDonald E, Wang R, LaVange LM, Pate V, Buse JB. Cancer incidence among those initiating insulin therapy with glargine versus human NPH insulin. Diabetes Care 2013;36:3517-25.

Active Comparator, New User Design

- Kramer et al. J Chron Dis 1987;40:1073-85:
  - ***"Compared with what?*** .. it is important to **compare** that risk **with** that of some **other real therapeutic option** for **patients with the same clinical indication**. Just as in a clinical trial investigating treatment efficacy, any epidemiologic study of treatment risks should **compare two or more viable treatment alternatives**."
  - ".. measuring risks **conditionally** on .. indication is .. **essential** to reduce confounding"
  - ***"For what period of time?*** The risk posed by a drug for a .. event **is not generally** the same in the **sixth month of chronic therapy** as in the **first or second week**."

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

New user part previously mentioned by Feinstein 1971 – see: Lund JL, Richardson DB, Stürmer T. The active comparator, new user study design in pharmacoepidemiology: historical foundations and contemporary application. Current Epidemiology Reports 2015;2:221-8.

Conclusions: Where Are We Now?

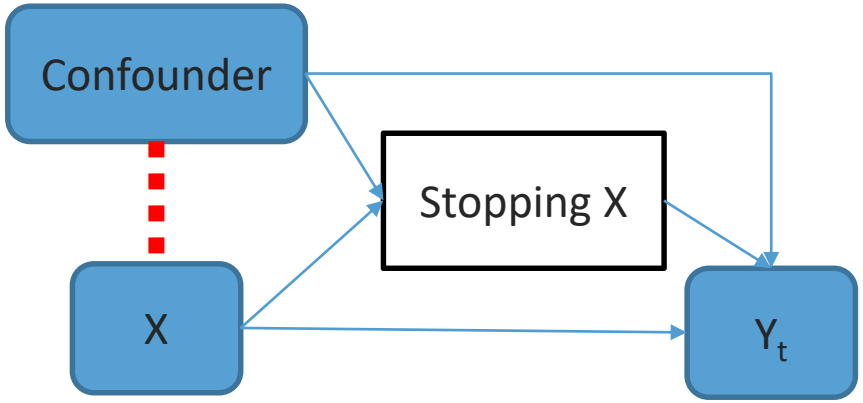
- Active comparator, new user design dramatically reduces potential for bias due to
  - Confounding by indication
  - Confounding by frailty
  - Non-adherence/time-varying hazards
  - Immortal time
- Focus on intervention needed for causal inference
- Comparator selection obviously important
- Standard design for nonexperimental CER

Where Are We Going?

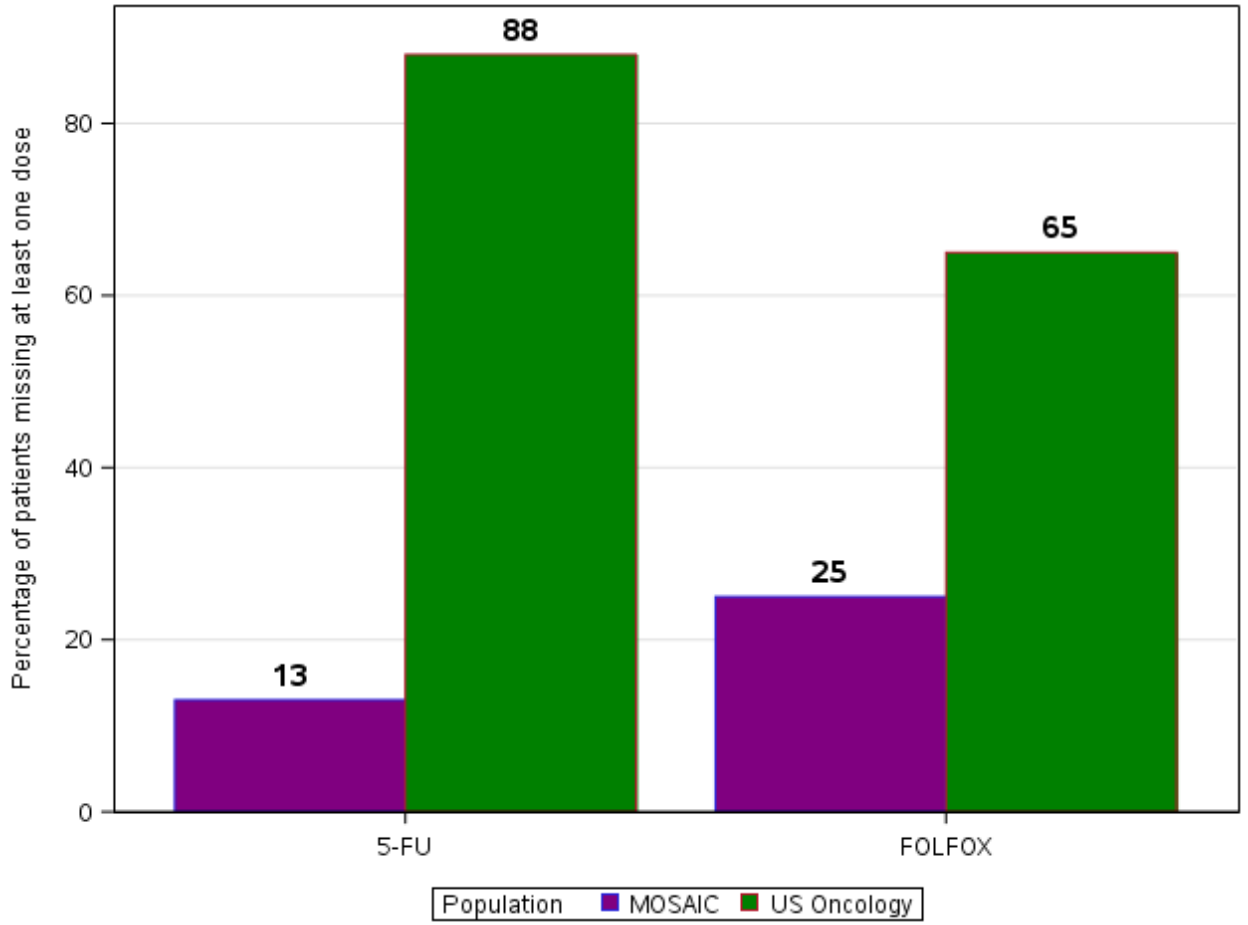
1. On-Treatment Estimates and Selection Bias

- If stopping study medication is differential by treatment and staying on treatment is affected by confounders, conditioning on remaining on treatment opens up a biasing path
- This path can be closed by *inverse probability of censoring weights*

PS: note that this is true in absence of baseline confounding, i.e., including RCTs!



Frequency Missing at Least One Chemo Dose

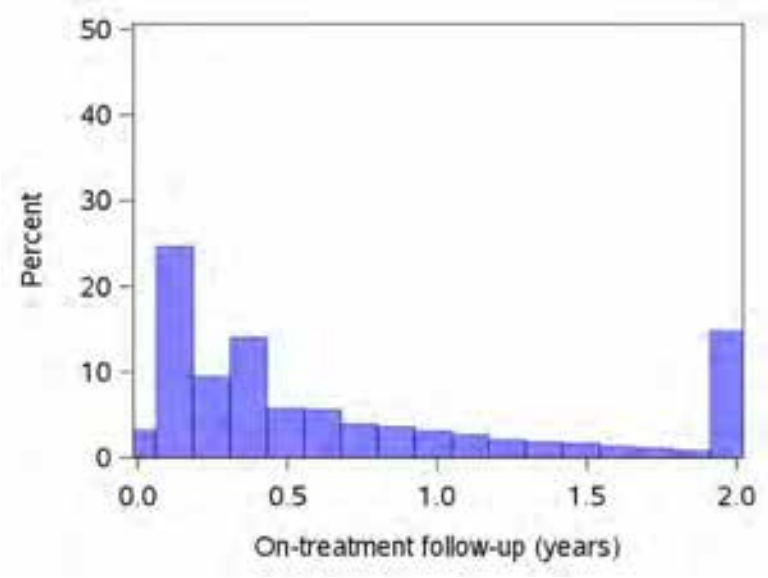


Jennifer L. Lund, PhD (PI)
Enhancing Hybrid Study Designs for CER
PCORI ME-2017C3-9337

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

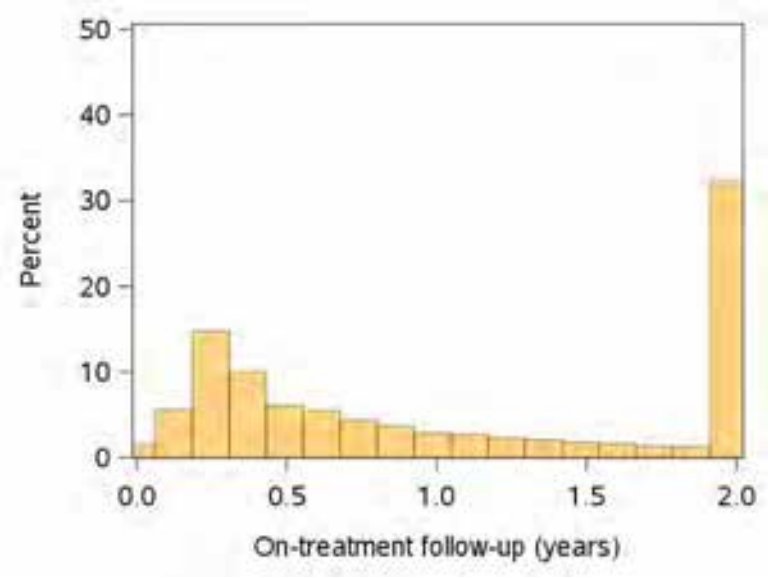
On-Treatment Follow-Up in US Medicare

Dabigatran New Users



Median OT FU: 152 days (P25: 60, P75: 382)

Warfarin New Users

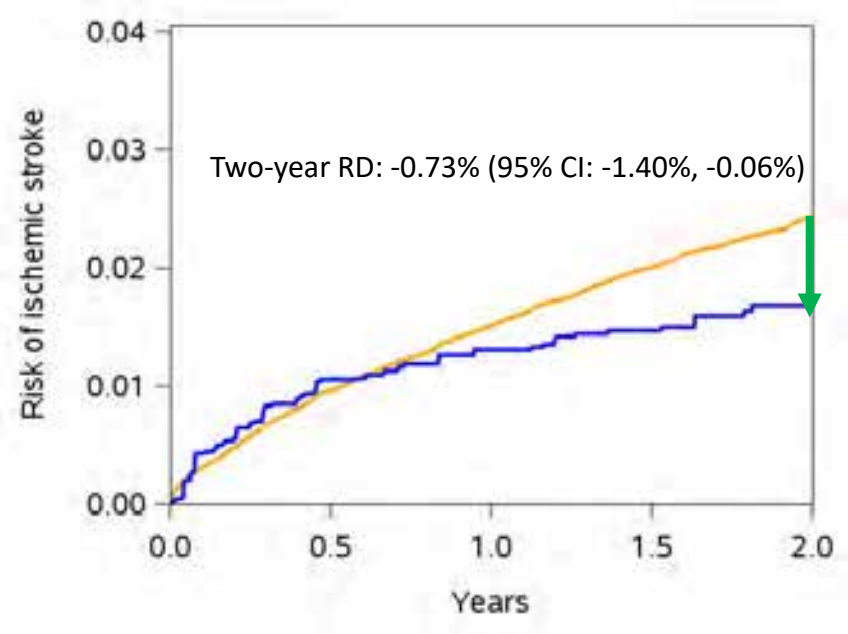


Median OT FU: 259 days (P25: 117, P75: 625)

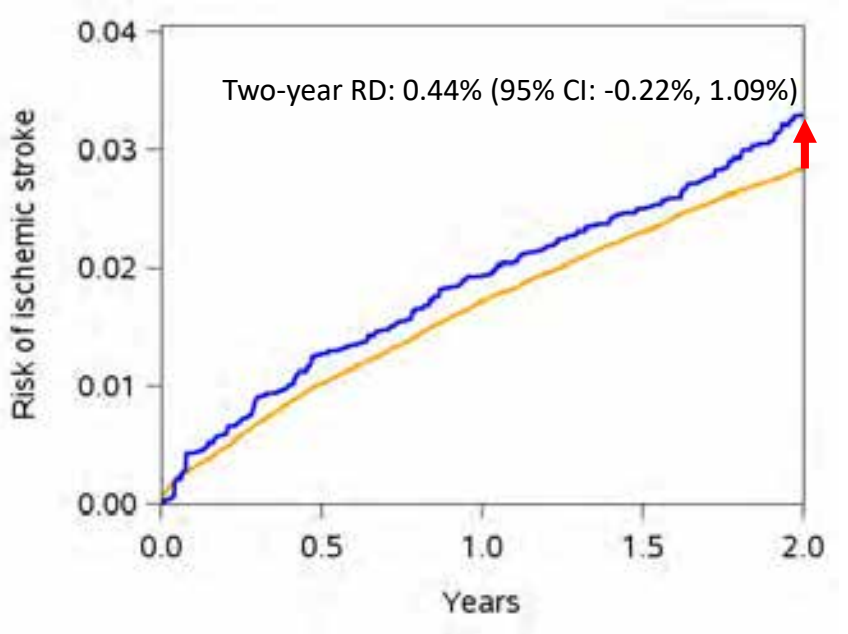
Slides Adapted from Michael Webster-Clark, PharmD, PhD, presented at 35th ICPE, Philadelphia, August 2019

Dabigatran vs Warfarin and Ischemic Stroke

On Treatment



Initial Treatment



Dabigatran new users

Warfarin new users

Conclusions On-Treatment Estimates

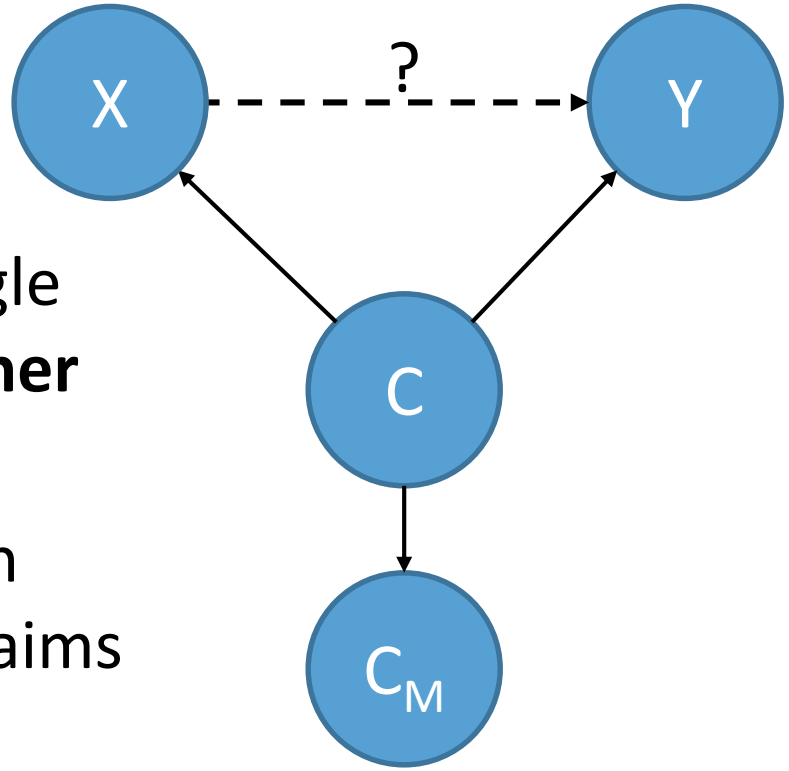
- The benefit (and harm) of treatments may not be realized in the real world due to lack of adherence
- This complicates RCT generalizability
- Methods to “account” for non-adherence depend on measured predictors of non-adherence
- Linkage of claims with e.g., EHR data will help with prediction
- Identification of barriers to adherence (subgroups most likely to benefit from interventions) important

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

Webster-Clark M et al., under external review

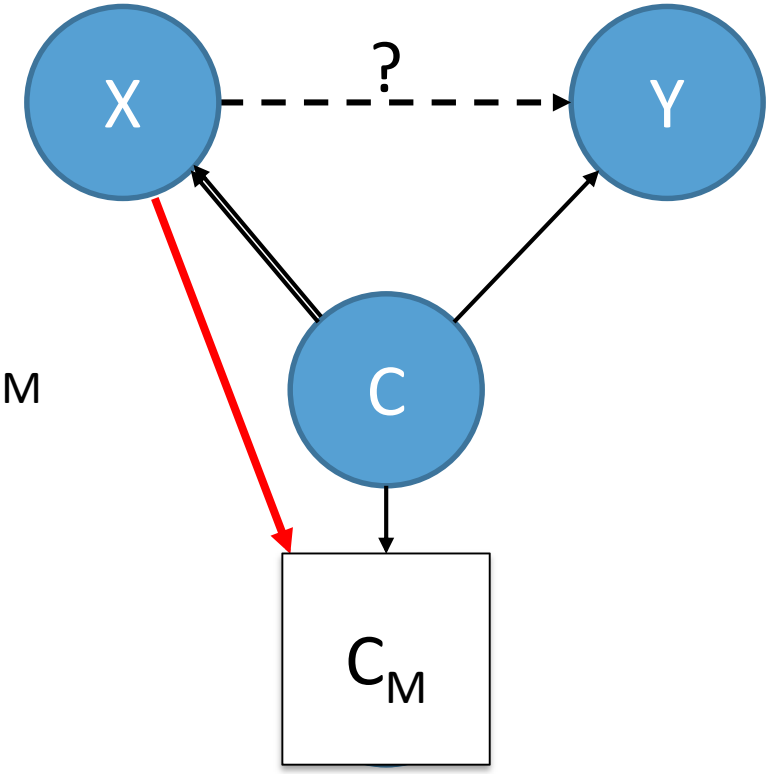
2. Single-Arm Trials and Confounder Adjustment

- We have data from a single arm trial of a preventive drug, as well as insurance claims (comparator)
- Physicians preferentially recruit patients that smoke **(C)** in the single arm trial since smokers are at **higher** risk for **(Y)**
- We can only measure C_M with high specificity but low sensitivity in claims



Three Major Graphical Conclusions

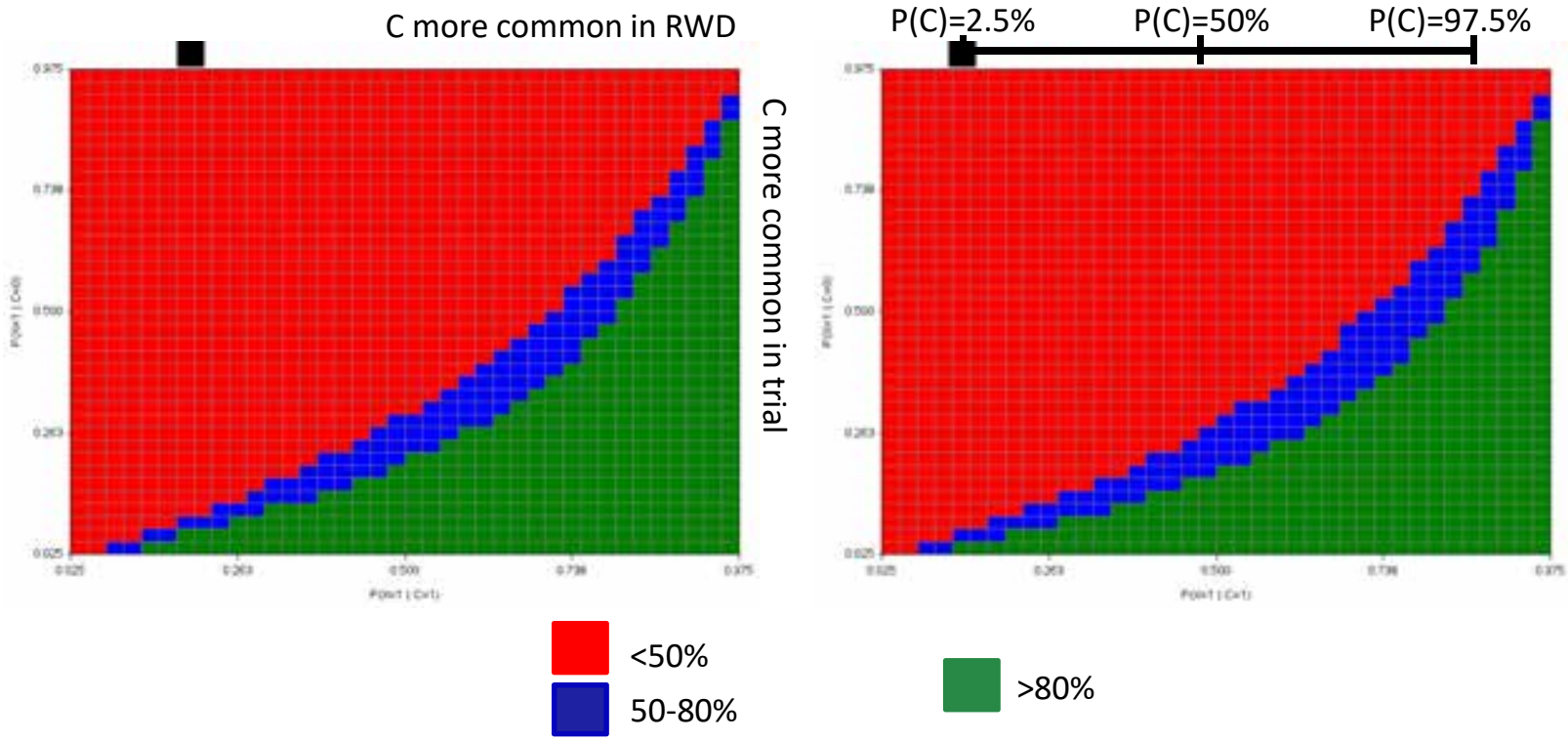
- If there is **no X->C_M arrow,** adjusting for C_M cannot generate bias
  - Will partially control for C
- If there is a **X->C_M arrow but no C->X arrow,** adjusting for C_M **will always** generate bias
- If both arrows exist, their direction and strength determine overall bias



Confounding Control When Sensitivity is Low (claims)

3-D Figure: P(C) = 10%

4-D Figure: P(C) ranges from 2.5% to 97.5%

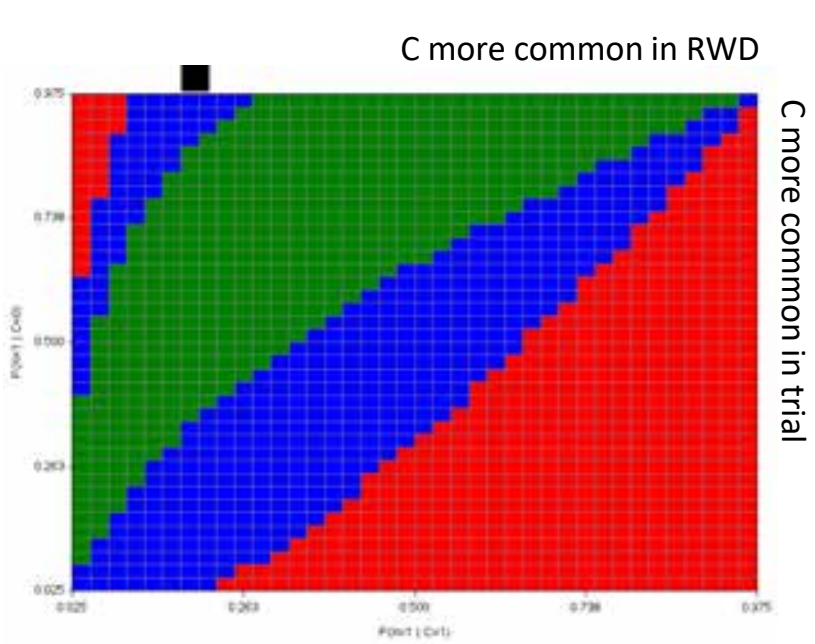


THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

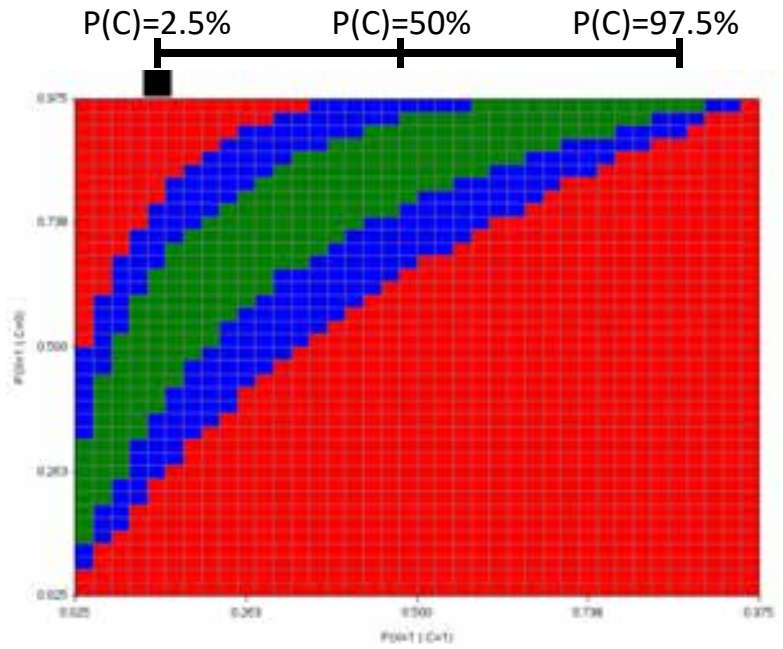
Michael Webster-Clark, PharmD, PhD, unpublished

Bias in Stratum C=1 When Specificity is 0.99

3-D Figure: P(C) = 10%



4-D Figure: P(C) ranges from 2.5% to 97.5%



>50%

20-50%

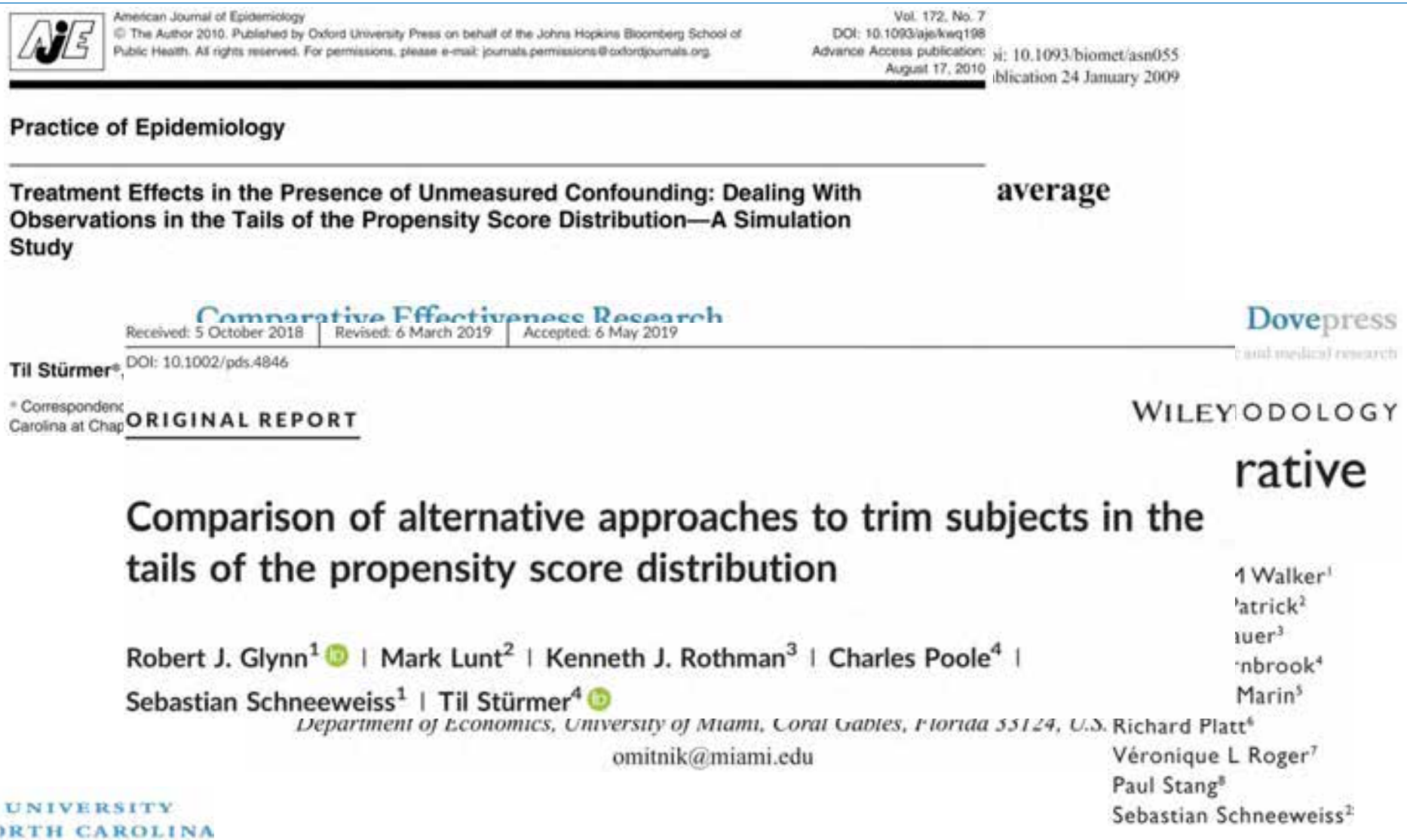
<20%

Conclusions: Single-Arm Trials and Confounder Adjustment

- If sensitivity or specificity of a covariate **differ**, the effects of controlling for C_M depend on:
  - Strength and direction of causal effects on X and Y
  - Prevalence of the confounder
  - Type and degree of differential misclassification
- Restriction is not always a solution, even when both data sets have high (not: perfect!) specificity
- We can identify parameter spaces where confounding can be sufficiently controlled for

Webster-Clark M et al., under internal review

3. PS to Identify Study Population at Equipoise



American Journal of Epidemiology

Vol. 172, No. 7
DOI: 10.1093/aje/kwq198
Advance Access publication: August 17, 2010

Practice of Epidemiology

Treatment Effects in the Presence of Unmeasured Confounding: Dealing With Observations in the Tails of the Propensity Score Distribution—A Simulation Study

average

Comparative Effectiveness Research
Received: 5 October 2018 | Revised: 6 March 2019 | Accepted: 6 May 2019
DOI: 10.1002/pds.4846

Dovepress

Til Stürmer*,

* Correspondenc
Carolina at Chap

ORIGINAL REPORT

WILEY ODOLOGY

rative

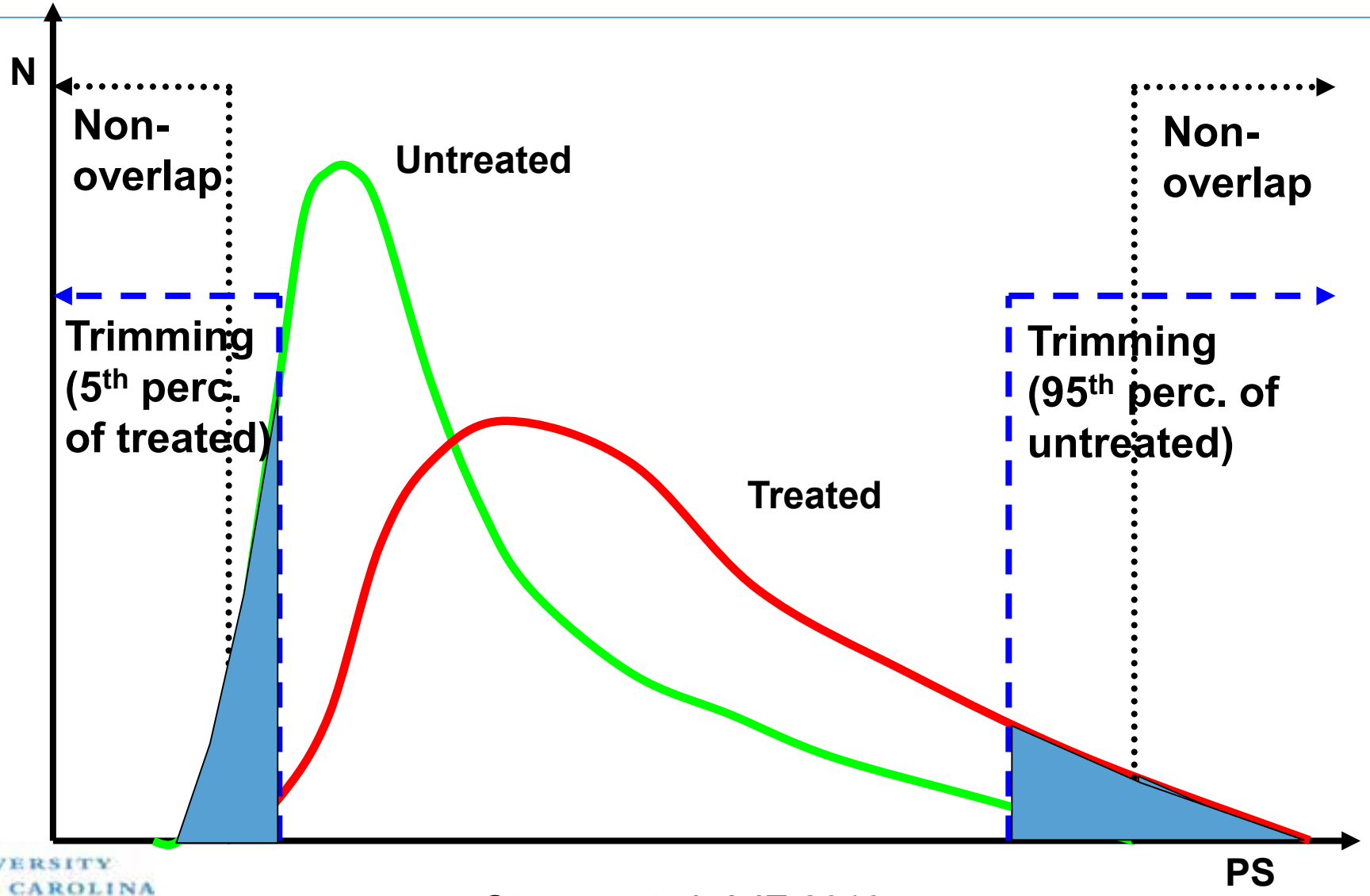
Comparison of alternative approaches to trim subjects in the tails of the propensity score distribution

Robert J. Glynn[1] | Mark Lunt[2] | Kenneth J. Rothman[3] | Charles Poole[4] | Sebastian Schneeweiss[1] | Til Stürmer[4]

Department of Economics, University of Miami, Coral Gables, Florida 33124, U.S.
omitnik@miami.edu

1 Walker[1]
atrick[2]
uer[3]
nbrook[4]
Marin[5]
Richard Platt[6]
Véronique L Roger[7]
Paul Stang[8]
Sebastian Schneeweiss[2]

Trimming Patients Treated Contrary to Prediction to Reduce Unmeasured Confounding by Frailty



Stürmer et al. AJE 2010

Conclusions: PS to Identify Study Population at Equipoise

- Focus on treatment decision is unique PS advantage
- Little equipoise between treatments in tails of PS
- Heterogeneity of treatment effects in tails plausibly due to unmeasured confounding (vs. real)
- Trimming small proportions of study population in tails of PS can improve validity
- Need more work/guidance on amount of trimming
- Define bias vs. treatment effect in target population (target validity; Westreich et al AJE 2019) promising

Til Stürmer (PI)
Propensity Scores and Preventive Drug Use in the Elderly.
National Institute on Aging (R01 AG056479)

4. Evaluate Adequacy of Sensitivity Analyses

Objectives

- Aid CDER in the development of guidance on use of sensitivity analyses to evaluate uncontrolled confounding
- Enhance the capacity of FDA to critically evaluate adequacy of sensitivity analyses of unmeasured confounding for assessment of non-experimental studies

Activities

- Identify and evaluate methods for assessment of bias due to uncontrolled confounding
- Provide tailored training for scientists at FDA
- Disseminate findings to research community

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

Michele Jonsson Funk (PI) FDA Contract Award No. 75F40119C10115. Methodological Advances in the Assessment of Uncontrolled Confounding

Thank you

sturmer@unc.edu

til.sturmer@post.harvard.edu

Session IV: Methodological and Analytical Considerations for Observational Studies

RCT replication with observational data

William Crown, PhD
Chief Scientific Officer, OptumLabs

October 3, 2019

OPTUM Labs®

Current (limited) literature suggests observational studies yield results similar to RCTs

The Cochrane Collaborative[1] examined 14 prior reviews comparing RCTs to observational studies:

Collectively, these reviews included data on 1,583 meta analyses spanning 228 medical conditions.

- 11 of 14 studies (79%) found no difference in ratios of odds ratios (ROR)
- One review suggested larger ROR for observational studies
- Two reviews suggested smaller ROR for observational studies

Earlier studies showed similar results.[2,3]

"Our results showed that, on average, there is little difference between the results obtained for RCTs and observational studies."

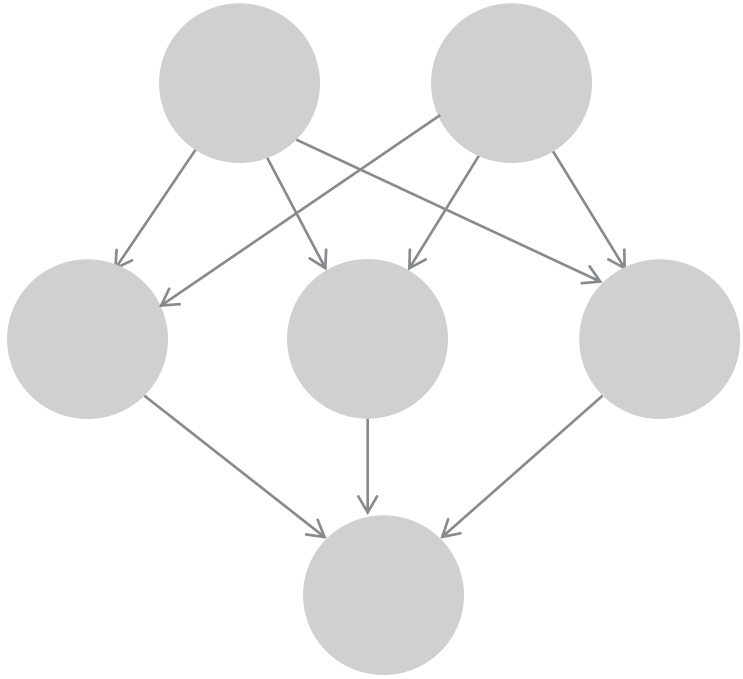
1. Anglemyer A, Horvath HT, Bero L. Healthcare outcomes assessed with observational study designs compared with those assessed in randomized trials (Review). The Cochrane Library 2014, Issue 4. 2. Benson K, Hartz AJ. A Comparison of Observation Studies and Randomized, Controlled Trials. N Engl J Med 2000; 342: 1878–86 . 3. Concato J, Shah N, Horwitz RI. Randomized, Controlled Trials, Observation Studies and the Hierarchy of Research Designs. N Engl J Med 2000; 342: 1887–92.

Causal frameworks are needed to actually replicate the RCTs

- Pearl, J (2013). Causality: Models, Reasoning, and Inference. 2nd Edition. New York, NY: Cambridge University Press.
- Van der Laan MJ, Rose S (2011). Targeted Learning: Causal Inference for Observational and Experimental Data. New York, NY: Springer-Verlag.
- Rubin, D (1974). Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies. Journal of Educational Psychology 64, 688-701.
- Heckman, J (1976). "The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variables and an Estimator for Such Models." Annals of Economic and Social Measurement 5: 475–492.
- Zellner A, Theil H (1962). Three-Stage Least Squares: Simultaneous Estimation of Simultaneous Equations." Econometrica 30(1):54-78.

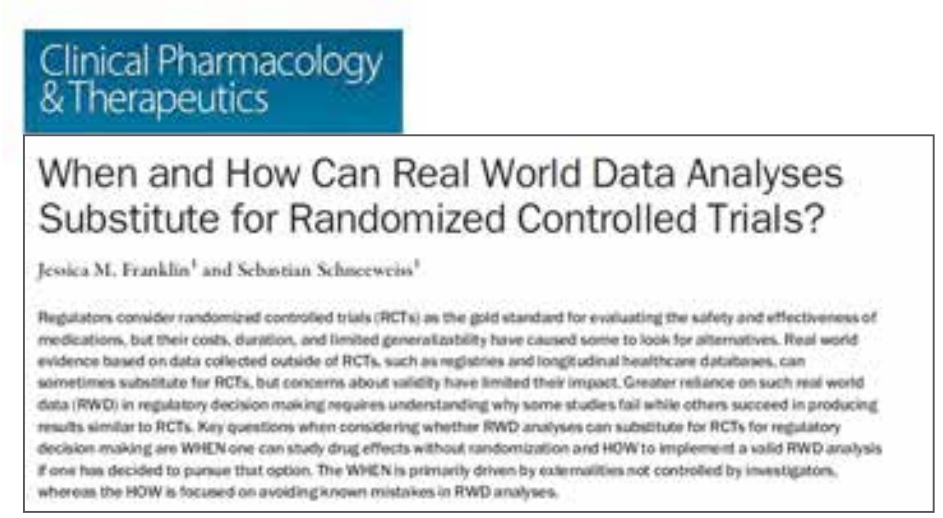
There are many methods for causal modeling with health care data

- Standard regression models with quasi-experimental design
- Propensity score matching or inverse probability weighting
- G estimation and marginal structural models
- Doubly robust methods
- Instrumental variables
- Differences in differences
- Targeted maximum likelihood estimation



We've learned a lot about how to do comparisons correctly

1. Active comparator, same treatment modality
2. New users
3. High-dimensional proxy adjustment
4. Control for medication adherence
5. Avoiding design flaws:
    a. reverse causation
    b. adjustment for causal intermediaries
    c. immortal time bias
    d. depletion of susceptibles


Clinical Pharmacology & Therapeutics

When and How Can Real World Data Analyses Substitute for Randomized Controlled Trials?

Jessica M. Franklin[1] and Sebastian Schneeweiss[1]

Regulators consider randomized controlled trials (RCTs) as the gold standard for evaluating the safety and effectiveness of medications, but their costs, duration, and limited generalizability have caused some to look for alternatives. Real world evidence based on data collected outside of RCTs, such as registries and longitudinal healthcare databases, can sometimes substitute for RCTs, but concerns about validity have limited their impact. Greater reliance on such real world data (RWD) in regulatory decision making requires understanding why some studies fail while others succeed in producing results similar to RCTs. Key questions when considering whether RWD analyses can substitute for RCTs for regulatory decision making are WHEN one can study drug effects without randomization and HOW to implement a valid RWD analysis if one has decided to pursue that option. The WHEN is primarily driven by externalities not controlled by investigators, whereas the HOW is focused on avoiding known mistakes in RWD analyses.

Franklin J. and Schneeweiss S. When and How Can Real World Data Analyses Substitute for Randomized Controlled Trials? Clinical Pharmacology and Therapeutics 2017.

There are a limited and growing number of observational studies replicating RCTs

Observational study followed by RCT:

- Schneeweiss S, Seeger J, Landon J, Walker A. Aprotinin during Coronary-Artery Bypass Grafting and Risk of Death. *NEJM* 358(8), 2008
- Fergusson D, Hebert P, Mazer D, et al. A Comparison of Aprotinin and Lysine Analogues in High-Risk Cardiac Surgery. *NEJM* 358(22), 2008

RCT followed by observational study:

- Connolly S, Ezekowitz M. Yusef S, et al. NEJM. Dabigatran versus Warfarin in Patients with Atrial Fibrillation. 361(12), 2009
- Seeger J, Bykov K, Bartels D, et al. Safety and Effectiveness of Dabigatran and Warfarin in Routine Care of Patients with Atrial Fibrillation. *Thrombosis and Haemostasis* 114(12):1277-89, 2015

Observational study conducted concurrently with RCT:

- Noseworthy PA, Gersh BJ, Kent DM, et al. Atrial fibrillation ablation in practice: Assessing CABANA generalizability. Eur Heart J. 2019 April 21;ehz085.

A high profile case where RCTs and observational studies differed

The Nurses Health Study (observational) had found a protective cardiovascular risk from HRT.

Stampfer MJ et al. Postmenopausal Estrogen Therapy and Cardiovascular Disease: Ten-Year Follow-up from the Nurses' Health Study. N. Engl. J. Med 325, 756-762 (1991).

The Women's Health Initiative (RCT) found just the opposite.

Rossouw JE et al. Risks and Benefits of Estrogen Plus Progestin in Healthy Postmenopausal Women: Principal Results from the Women's Health Initiative Randomized Controlled Trial. JAMA 288, 321-333 (2002)

Was randomization the issue?

And subsequent studies revealed the reasons why.

Hernan MA et al. Observational Studies Analyzed Like Randomized Experiments: An Application to Postmenopausal Hormone Therapy and Coronary Heart Disease. Epidemiology 19, 766-779 (2008)

Goodman SN, Schneeweiss S. and Baiocchi M. Using Design Thinking to Differentiate Useful From Misleading Evidence in Observational Research. JAMA 317, 705-707 (2017).

Study design was the difference.

What is the role of real-world data in regulatory decision making?

OPERAND (Observational Patient Evidence for Regulatory Approval and uNderstanding Disease)

Improve the confidence in observational data to generate evidence supporting treatment effectiveness and safety for patient populations beyond those studied in randomized clinical trials (RCTs).

Approach

- Replicate two clinical trials: ROCKET for atrial fibrillation and Lead-2 for Type 2 diabetes control
  - Using OLDW claims and clinical data
  - Applying methods expertise
- Engage diverse experts in government, academia, industry to advise the program

Potential impact

- Inform policy on the use of real-world evidence to support regulatory approvals of new drug indications and to satisfy post-approval safety surveillance requirements
- Validation of using observational data to complement evidence from RCTs
- Innovation in clinical trial design, thereby bringing new treatments to market faster and more cost-effectively

Co-leads

OPTUM Labs

MULTI-REGIONAL CLINICAL TRIALS

Sponsors

Amgen | AstraZeneca | Merck | Optum | Pfizer | Sanofi | UCB BioSciences, Inc.

Research partners

BROWN School of Public Health

DEPARTMENT OF POPULATION MEDICINE HARVARD MEDICAL SCHOOL Harvard Pilgrim Health Care Institute

Expert panel

Duke-Margolis Center for Health Policy

Eli Lilly & Company

GlaxoSmithKline

Food and Drug Administration

ISPOR

National Pharmaceutical Council

…and more

OPERAND study design

Focus: On-label effectiveness in defined subgroups

Number of teams and trials	Two academic institutions will independently replicate two identical target trials: 1. ROCKET for atrial fibrillation 2. Lead-2 for Type 2 diabetes control
Data	• (a) Claims data alone and (b) Claims + EHR, each used for sensitivity analyses • Data will be restricted to inclusion and exclusion criteria of pivotal RCT and on-label indication
Methodology	Bootstrapping methods along with bias analysis will be used to understand variability in treatment effect estimates
Documentation	Research team must document assumptions and choices made when emulating trials
Approach	To ensure comparability, the teams will: • Be given a common clinical question and the study RCT protocol • Be given defined set of anticipated methods • Have flexibility to use their own methods in certain areas • Initially, be restricted to inclusion/exclusion criteria

Measures of replication

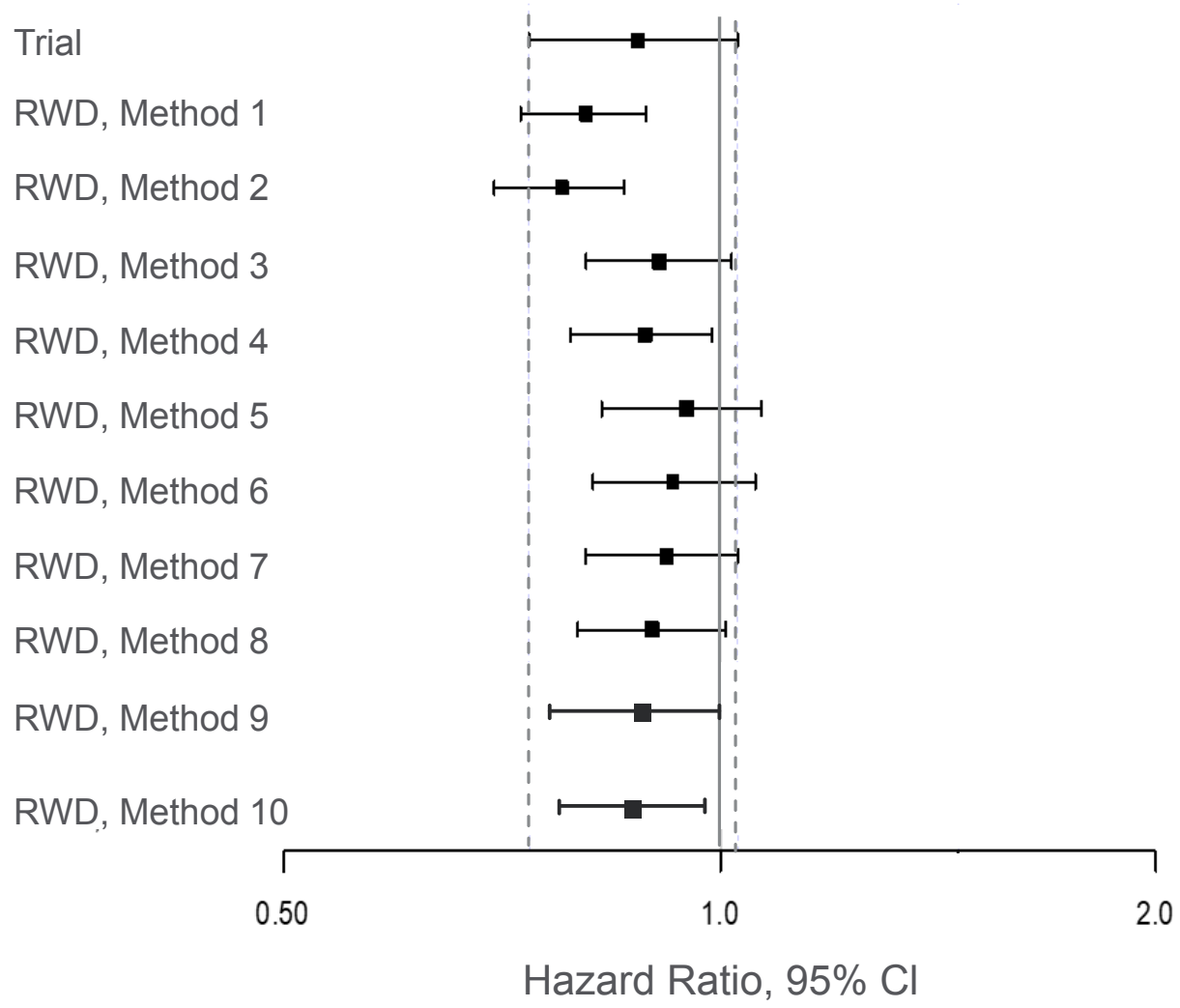
Regulatory agreement

Defined as statistically significant result with directional equivalence between the RCT and observational study.

Estimate agreement

Defined as the point estimate of the observational study falling within the 95% confidence interval of the ATE from the RCT using the reported standard errors of the RCT to define the confidence interval.

Preliminary Results: Distribution of estimates from ROCKET AF Trial and the replication study



The potential for using supervised machine learning methods

Traditionally machine learning methods focused on prediction and classification — not causal inference

Many methods	
• Classification trees • Random forests • Bagging and boosting models • Ridge, lasso, and elastic net regression	• Support vector machines • Ensembles • Neural networks • And many others…

Hastie T., Tibshirani R., Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd Edition. New York: Springer.

Is causal inference compatible with machine learning?

There are two paths forward:

1) Sequential approach
   - Estimate prediction/classification models using machine learning techniques to select features
   - Estimate causal models with epidemiologic or econometric approaches using selected features in the model specifications

2) Targeted Maximum Likelihood Estimation (TMLE)

A snapshot of targeted maximum likelihood estimation

American Journal of Epidemiology

American Journal of Epidemiology

Vol. 185, No. 1
DOI: 10.1093/aje/kww165
Advance Access publication:
December 9, 2016

Practice of Epidemiology

Targeted Maximum Likelihood Estimation for Causal Inference in Observational Studies

Megan S. Schuler and Sherri Rose*

* Correspondence to Dr. Sherri Rose, Department of Health Care Policy, Harvard Medical School, 180 Longwood Avenue, Boston, MA 02215 (e-mail: rose@hcp.med.harvard.edu).

Initially submitted August 19, 2015; accepted for publication November 1, 2016.

Estimation of causal effects using observational data continues to grow in popularity in the epidemiologic literature. While many applications of causal effect estimation use propensity score methods or G-computation, targeted maximum likelihood estimation (TMLE) is a well-established alternative method with desirable statistical properties. TMLE is a doubly robust maximum-likelihood–based approach that includes a secondary "targeting" step that optimizes the bias-variance tradeoff for the target parameter. Under standard causal assumptions, estimates can be interpreted as causal effects. Because TMLE has not been as widely implemented in epidemiologic research, we aim to provide an accessible presentation of TMLE for applied researchers. We give step-by-step instructions for using TMLE to estimate the average treatment effect in the context of an observational study. We discuss conceptual similarities and differences between TMLE and 2 common estimation approaches (G-computation and inverse probability weighting) and present findings on their relative performance using simulated data. Our simulation study compares methods under parametric regression misspecification; our results highlight TMLE's property of double robustness. Additionally, we discuss best practices for TMLE implementation, particularly the use of ensembled machine learning algorithms. Our simulation study demonstrates all methods using super learning, highlighting that incorporation of machine learning may outperform parametric regression in observational data settings.

causal inference; machine learning; observational studies; super learner; targeted maximum likelihood estimation

Questions?

Session IV: Methodological and Analytical Considerations for Observational Studies

Real World Evidence from Healthcare Databases: We have come a long way

Sebastian Schneeweiss, MD, ScD
Professor of Medicine and Epidemiology

Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine
Brigham and Women's Hospital, Harvard Medical School, Boston

October 2019

Funding

- This study was funded in part by FDA HHSF223201710186C and HHSF…46C
- This study was funded in part by the NHLBI
- Additional funding came from PCORI

Disclosures

- PI, Harvard-Brigham & Women's Hospital Drug Safety Research Center (FDA)
- Co-Chair, Methods Core of the FDA Sentinel System
- Co-Chair, Partners Center for Integrated Healthcare Data Research
- PI of research grants awarded to BWH by Bayer, Vertex, Boehringer Ingelheim
- Consulting fees from WHISCON, LLC, and Aetion, Inc. (incl. equity)
- Grants/contracts from NIH, AHRQ, PCORI, FDA, IMI, Arnold Foundation

Reminder: Why we love RCTs

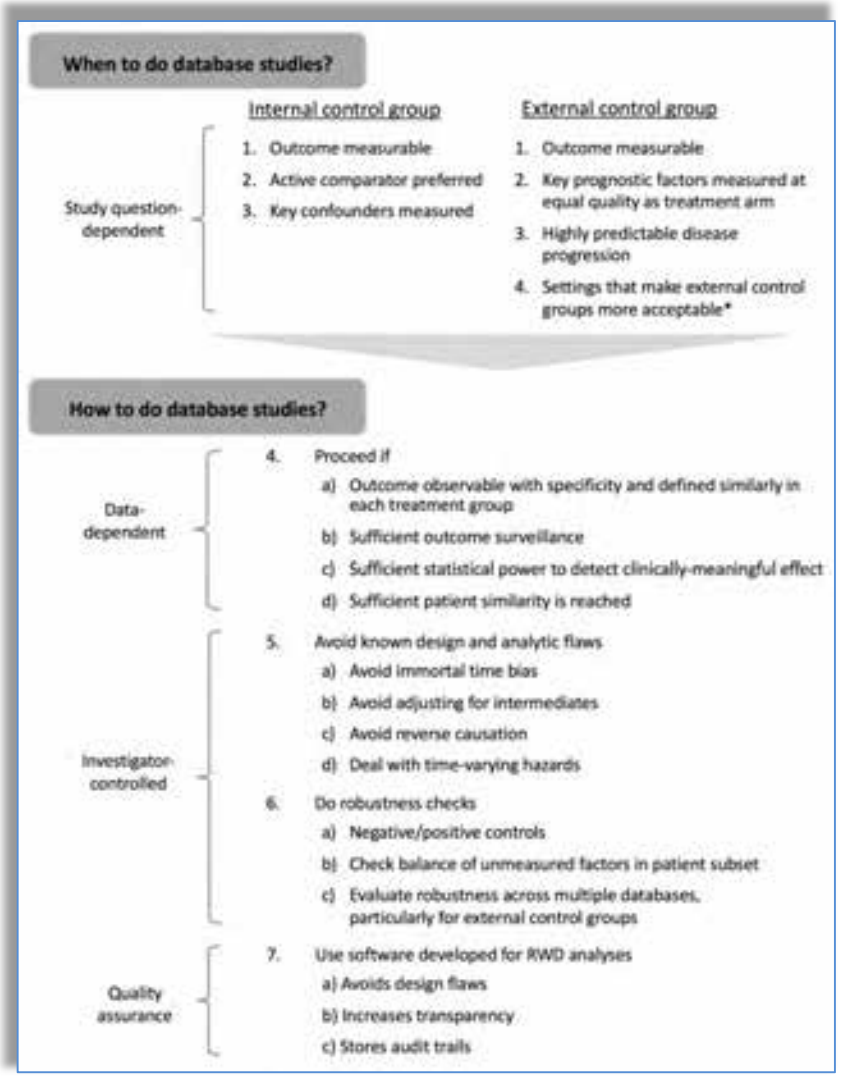
Randomized Controlled Trials

Random treatment assignment

Controlled outcome measurement

Easy to understand and communicate

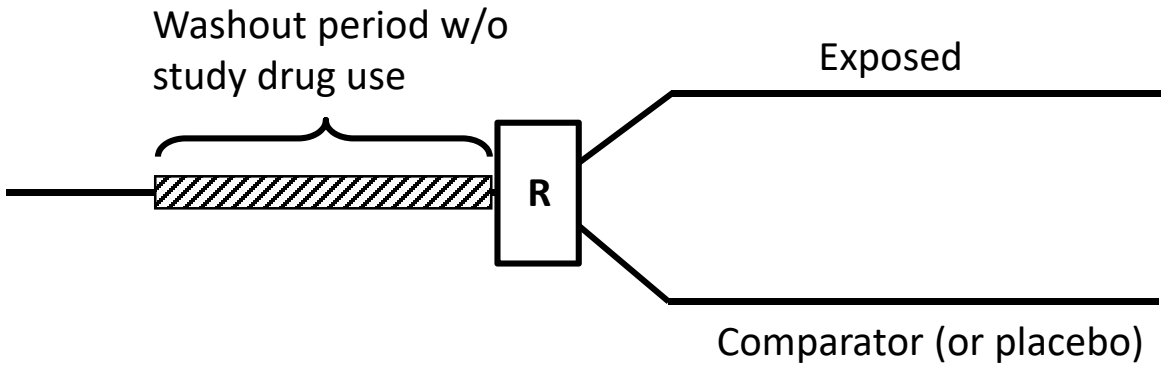
And how we get to causal associations with RWE



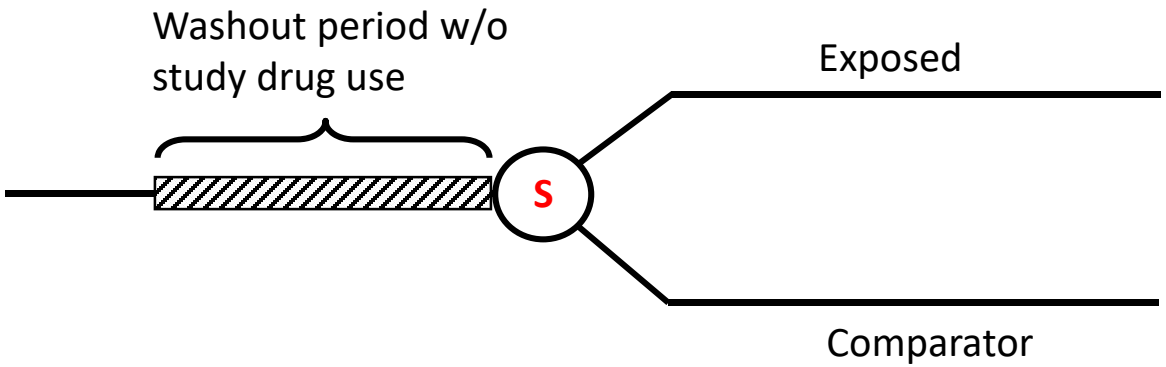
Franklin, Glynn, Martin, Schneeweiss, CP&T 2019

Causal study designs: Contemplate the target trial

Parallel group RCT



Cohort study



= New-user, active-comparator cohort study

1) Why do we like new user cohort studies?

- Patients at a clear inception point
- Confounders measured before exposure
- Compatible with propensity score analyses
- Allows to describe time-varying hazards
- Also reduces the risk of immortal time

2) Why do we like active comparators?

- Patients are more similar

RCT

RWE

New users

Current users

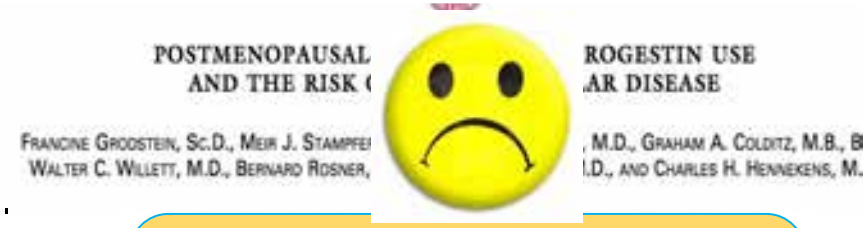

Estrogen plus Progestin and the Risk of Coronary Heart Disease

JoAnn E. Manson, M.D., Dr.P.H., Judith Hsia, M.D., Karen C. Johnson, M.D., M.P.H., Jacques E. Rossouw, M.D., Annlouise R. Assaf, Ph.D., Norman L. Lasser, M.D., Ph.D., Maurizio Trevisan, M.D., Henry R. Black, M.D., Susan R. Heckbert, M.D., Ph.D., Robert Detrano, M.D., Ph.D., Ora L. Strickland, Ph.D., Nathan D. Wong, Ph.D., John R. Crouse, M.D., Evan Stein, M.D., and Mary Cushman, M.D., for the Women's Health Initiative Investigators*

Risk of CHD events
1.68 (1.15–2.45)


Observatio d Like
Rand s
An Application to one Therapy
and C

Risk of CHD events
1.42 (0.92–2.20)


POSTMENOPAUSAL ROGESTIN USE
AND THE RISK AR DISEASE
FRANCINE GROODSTEIN, SC.D., MEIR J. STAMPFE M.D., GRAHAM A. COLDITZ, M.B., B
WALTER C. WILLETT, M.D., BERNARD ROSNER, D., AND CHARLES H. HENNEKENS, M.

Risk of CHD events
0.60 (0.43–0.83)


Dabigatran versus Warfarin in Patients with Atrial Fibrillation

Stuart J. Connolly, M.D., Michael D. Ezekowitz, M.B., Ch.B., D.Phil., Salim Yusuf, F.R.C.P.C., D.Phil., John Eikelboom, M.D., Jonas Oldgren, M.D., Ph.D., Amit Parekh, M.D., Janice Pogue, M.Sc., Paul A. Reilly, Ph.D., Ellison Themeles, B.A., Jeanne Varrone, M.D., Susan Wang, Ph.D., Marco Alings, M.D., Ph.D., Denis Xavier, M.D., Jun Zhu, M.D., Rafael Diaz, M.D., Basil S. Lewis, M.D., Harald Darius, M.D., Hans-Christoph Diener, M.D., Ph.D., Campbell D. Joyner, M.D., Lars Wallentin, M.D., Ph.D., and the RE-LY Steering Committee and Investigators*

Risk of stroke
0.66 (0.53–0.82)


ORIGINAL ARTICLE
Sequential Mo Comparative
Effectiveness a abigatran in
Routine Care

Risk of stroke
0.75 (0.58–0.98)


Dabigatran use in Danish atrial
fibrillation pat a
nationwide stu
Rikke Sørensen,[1,2] Gunnar G ersen,[4]
Jonas Bjerring Olesen,[1] Emil dt,[1] Deniz Karasoy,[1]
Morten Lamberts,[1] Mette Cha keke,[1] Gregory Y H Lip,[3]
Morten Lock Hansen[1]

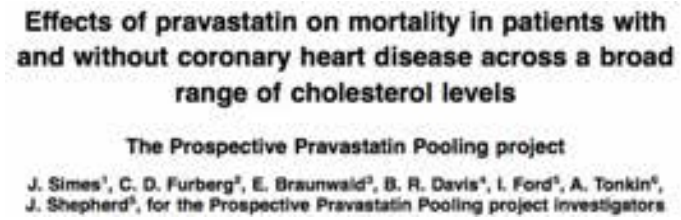
Risk of stroke
5.79 (1.81 to 18.6)

RCT

RW

Active comparator

Non-user comparator


Effects of pravastatin on mortality in patients with and without coronary heart disease across a broad range of cholesterol levels

The Prospective Pravastatin Pooling project

J. Simes[1], C. D. Furberg[2], E. Braunwald[3], B. R. Davis[4], I. Ford[5], A. Tonkin[6], J. Shepherd[7], for the Prospective Pravastatin Pooling project investigators

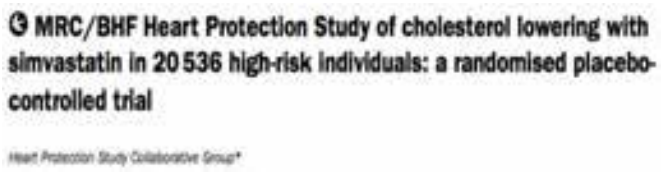
Risk of death (any)
0.78 (0.68–0.89)


Increasing Levels of ... Database Studies ... Comparison With ... Results

Risk of death (any)
0.79 (0.60–1.03)


Increasing Levels of Restriction in Pharmacoepidemiologic Database Studies ... Comparison With ... Results

Risk of death (any)
0.62 (0.58–0.66)


MRC/BHF Heart Protection Study of cholesterol lowering with simvastatin in 20 536 high-risk individuals: a randomised placebo-controlled trial

Heart Protection Study Collaborative Group*

Risk of hip fracture
1.05 (0.88–1.25)


Use of Statins ... Fractures

Risk of hip fracture
1.02 (0.83–1.24)


Inhibitors of hydroxymeth... reductase and risk of fracture among older w...

Risk of hip fracture
0.48 (0.27–0.83)

How to …

Data-dependent

4. Proceed if
   a) Outcome observable with specificity
   b) Sufficient outcome surveillance
   c) Sufficient patient similarity is reached[1)]

Investigator-controlled

5. Avoid known design and analytic flaws
   a) Avoid immortal time bias
   b) Avoid adjusting for causal intermediates
   c) Avoid reverse causation
   d) Deal with time-varying hazards

6. Do robustness checks
   a) Negative/positive controls
   b) Check balance of unmeasured factors

Franklin, Glynn, Martin, Schneeweiss, CPT 2019

RCT

Canagliflozin and Cardiovascular and Renal Events in Type 2 Diabetes

Bruce Neal, M.B., Ch.B., Ph.D., Vlado Perkovic, M.B., B.S., Ph.D., Kenneth W. Mahaffey, M.D., Dick de Zeeuw, M.D., Ph.D., Greg Fulcher, M.D., Ngozi Erondu, M.D., Ph.D., Wayne Shaw, D.S.L., Gordon Law, Ph.D., Mehul Desai, M.D., and David R. Matthews, D.Phil., B.M., B.Ch., for the CANVAS Program Collaborative Group*

Risk of death (any)
0.87 (0.74–1.01)

RWE

No immortal time

Use of sodium glucos… …bitors and risk of major cardiovascular … …re: Scandinavian register based cohort

Björn Pasternak, Peter Ueda, … …sson, Stefan Franzén, Soffia Gudbjörnsdottir, Kristia… Viktor Wintzell, Mads Melbye, Henrik Svanst…

Risk of death (any)
0.80 (0.69–0.92)

Immortal study time

Lower Risk of Hea… …Death in Patients Initiated on Sodiu… …ransporter-2 Inhibitors Versus … …Lowering Drugs

The CVD-REAL Study (Co… …ess of Cardiovascular Outcomes in New Users … …otransporter-2 Inhibitors)

Risk of death (any)
0.49 (0.41–0.57)

Transparency in process and implementation

Real-World Evidence of Treatment Effects: The Useful and the Misleading

Sebastian Schneeweiss[1,*]

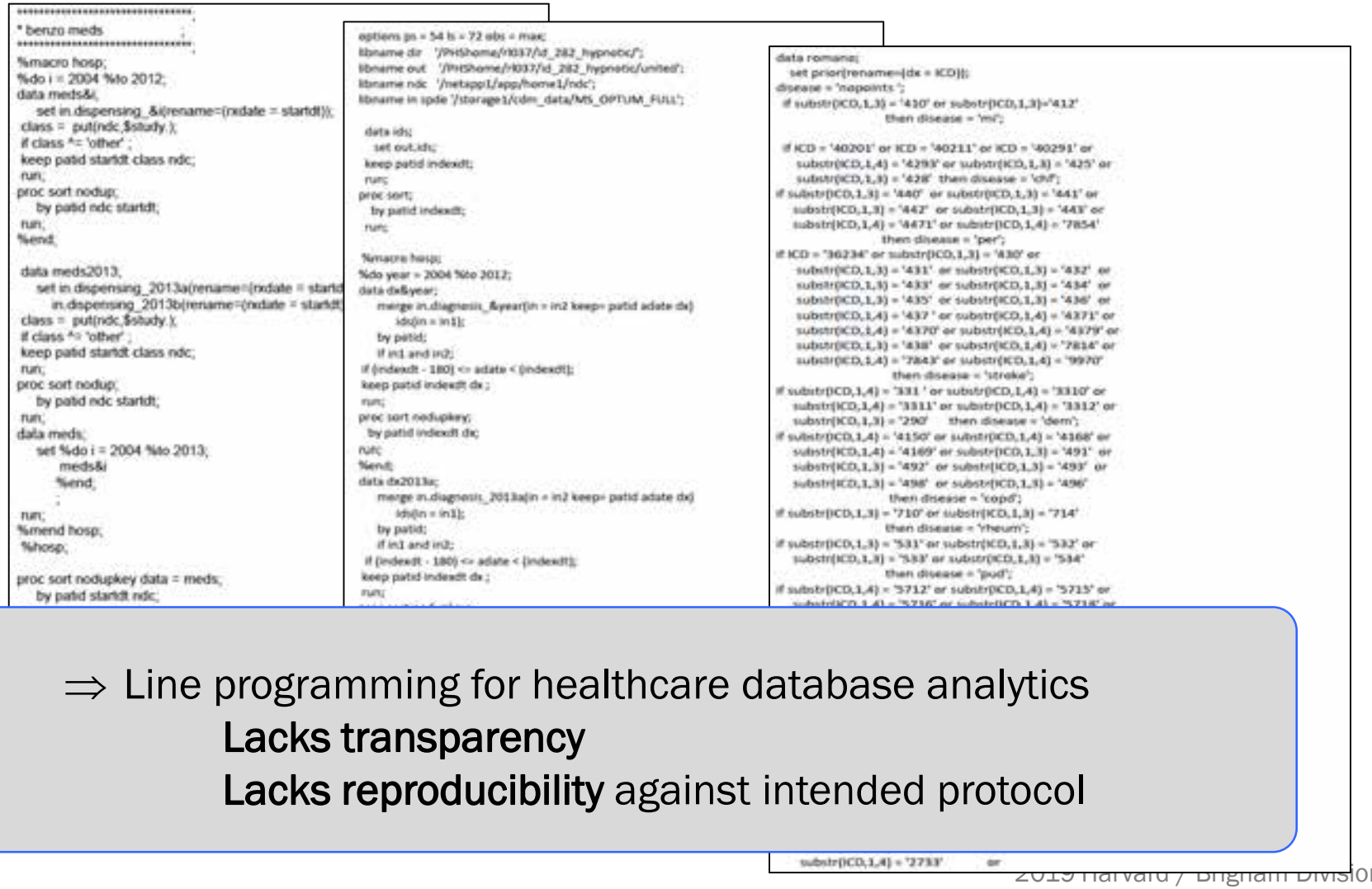
Turning real-world data (RWD) analyses into real-world evidence (RWE) requires accurate estimation of causal treatment effects; to convince its critics, advocates of RWE will need to get this right, reliably and predictably. This may sound like a high bar. However, if we are tired of seeing RWE categorically disregarded, we need to acknowledge that there are flawed RWD analyses and identify tools to quickly and confidently discriminate between actionable RWE and erroneous RWE.

- Transparency of implementation
  - Protocol + registration
- Reproducibility of implementation
- Validity/robustness of findings

Schneeweiss, CP&T 2019

Nothing wrong with sharing programming code but is not helpful…

… as it does not clarify whether the indented study was implemented accurately



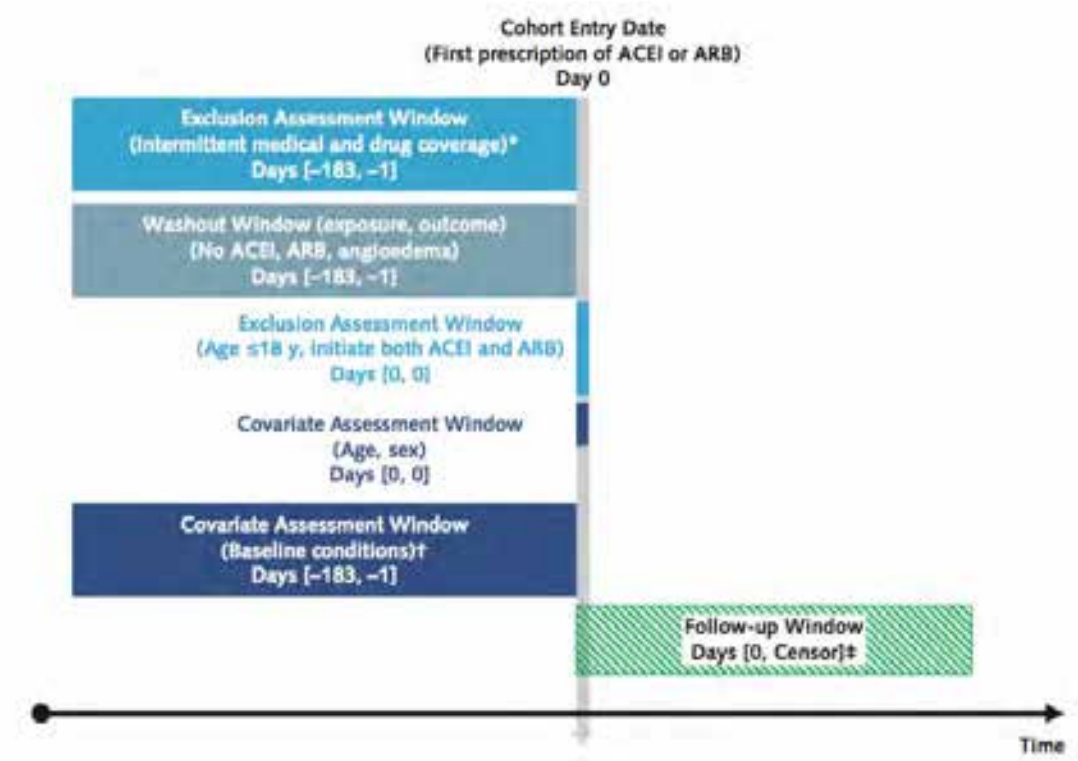
⇒ Line programming for healthcare database analytics

Lacks transparency

Lacks reproducibility against intended protocol

Making it easier for decision makers to fully understand RWE

Longitudinal design visualization




Annals of Internal Medicine RESEARCH AND REPORTING METHODS

Graphical Depiction of Longitudinal Study Designs in Health Care Databases

Sebastian Schneeweiss, MD, ScD; Jeremy A. Rassen, ScD; Jeffrey S. Brown, PhD; Kenneth J. Rothman, DrPH; Laura Happe, PharmD, MPH; Peter Arlett, MD; Gerald Dal Pan, MD, MHS; Wim Goettsch, PhD; William Murk, PhD; and Shirley V. Wang, PhD

Study design parameter table

Wang et al. in preparation with an FDA-HTA-industry consortium

Reporting to Improve Reproducibility and Facilitate Validity Assessment for Healthcare Database Studies V1.0

Shirley V. Wang[1,2] | Sebastian Schneeweiss[1,2] | Marc L. Berger[3] | Jeffrey Brown[4] | Frank de Vries[5] | Ian Douglas[6] | Joshua J. Gagne[1,2] | Rosa Gini[7] | Olaf Klungel[8] | C. Daniel Mullins[9] | Michael D. Nguyen[10] | Jeremy A. Rassen[11] | Liam Smeeth[6] | Miriam Sturkenboom[12] |

on behalf of the joint ISPE-ISPOR Special Task Force on Real World Evidence in Health Care Decision Making

ispe International Society for Pharmacoepidemiology

International Society for Pharmacoeconomics and Outcomes Research

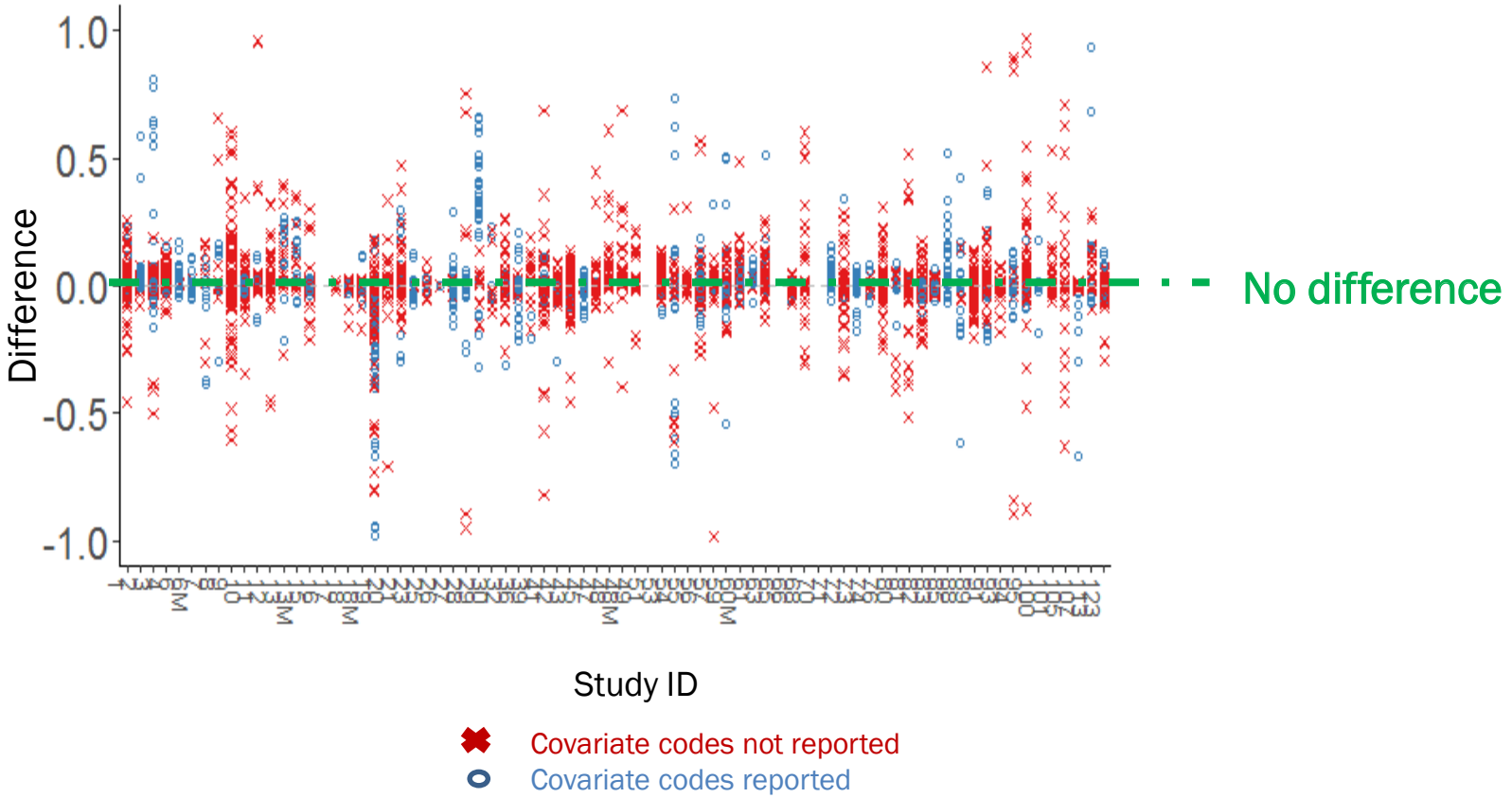
TABLE 2 Reporting specific parameters to increase reproducibility of database studies*

	Description	Example	Synonyms
A. Reporting on data source should include:			
A.1 Data provider	Data source name and name of organization that provided data.	Medicaid Analytic Extracts data covering 50	
A.2 Data extraction date (DED)	The date (or version number) wh extracted from the dynamic ra data stream (e.g. date that the for research use by the vendo		
A.3 Data sampling	The search/extraction criteria ap source data accessible to the r subset of the data available fro		
A.4 Source data range (SDR)	The calendar time range of data study. Note that the implemen use only a subset of the availa		
D. Reporting on exposure definition should include:			
D.1 Type of exposure	The type of exposure that is captured or measured, e.g. drug versus procedure, new use, incident, prevalent, cumulative, time-varying.	We evaluated risk of outcome Z following incident exposure to drug X or drug Y. Incident exposure was defined as beginning on the day of the first dispensation for one of these drugs after at least 180 days without dispensations for either (SED). Patients with incident exposure to both drug X and drug Y on the same SED were excluded. The exposure risk window for patients with Drug X and Drug Y began 10 days after incident exposure and continued until 14 days past the last days supply, including refills. If a patient refilled early, the date of the early refill and subsequent refills were adjusted so that the full days supply from the initial dispensation was counted before the days supply from the next dispensation was tallied. Gaps of less than or equal to 14 days in between one dispensation plus days supply and the next dispensation for the same drug were bridged (i.e. the time was	
D.2 Exposure risk window (ERW)	The ERW is specific to an exposure and the outcome under investigation. For drug exposures, it is equivalent to the time between the minimum and maximum hypothesized induction time following ingestion of the molecule.		Drug era, risk window
D.2a Induction period[1]	Days on or following study entry date during which an outcome would not be counted as "exposed time" or "comparator time".		Blackout period
D.2b Stockpiling[1]	The algorithm applied to handle leftover days supply if there are early refills.		
D.2c Bridging exposure episodes[1]	The algorithm applied to handle gaps that are longer than expected if there was perfect adherence (e.g. non-overlapping dispensation + day's supply).		Episode gap, grace period, persistence window, gap days

Replicating 150 database studies

Differences in binary/categorical characteristics* of cohort (publication – replication)

PI: Wang, Schneeweiss

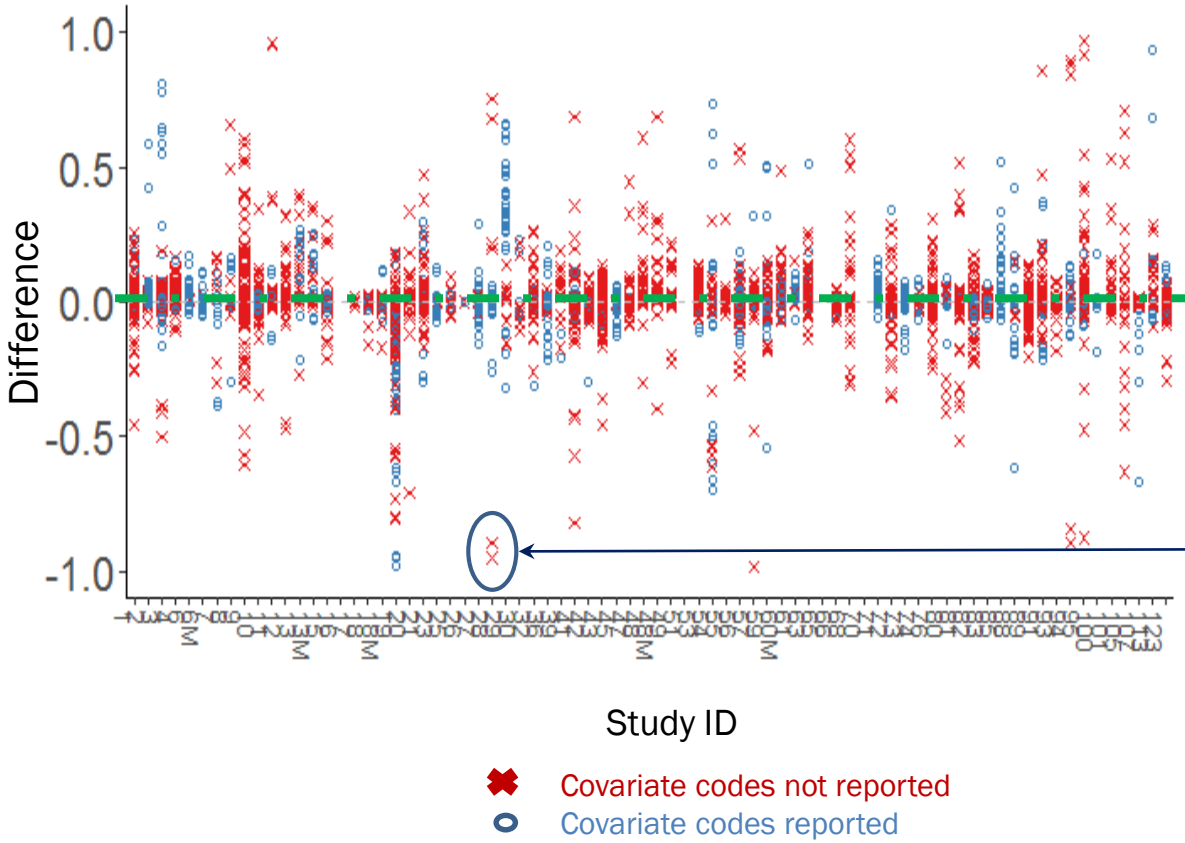


* binary/categorical

Replicating 150 database studies

Differences in binary/categorical characteristics* of cohort (publication – replication)

PI: Wang, Schneeweiss



- Authors provided citation to comorbidity score
- All patients in replication had score ≥ 2 because tumor/malignancy was part of inclusion
- > 75% in original had score = 0

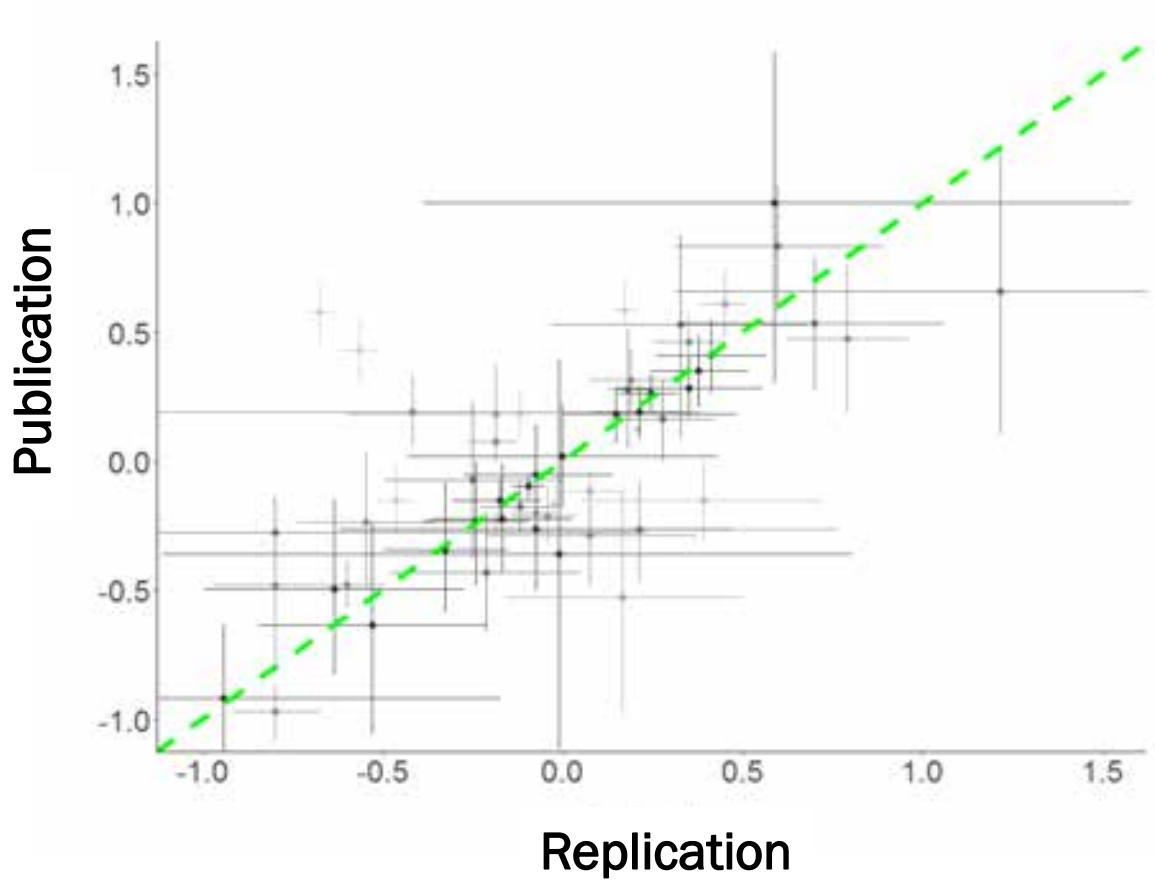
* binary/categorical

Effect Size and Confidence Limits

PI: Wang, Schneeweiss

Correlation between effect sizes
(publication vs replication)

- Correlation coefficient = 0.74

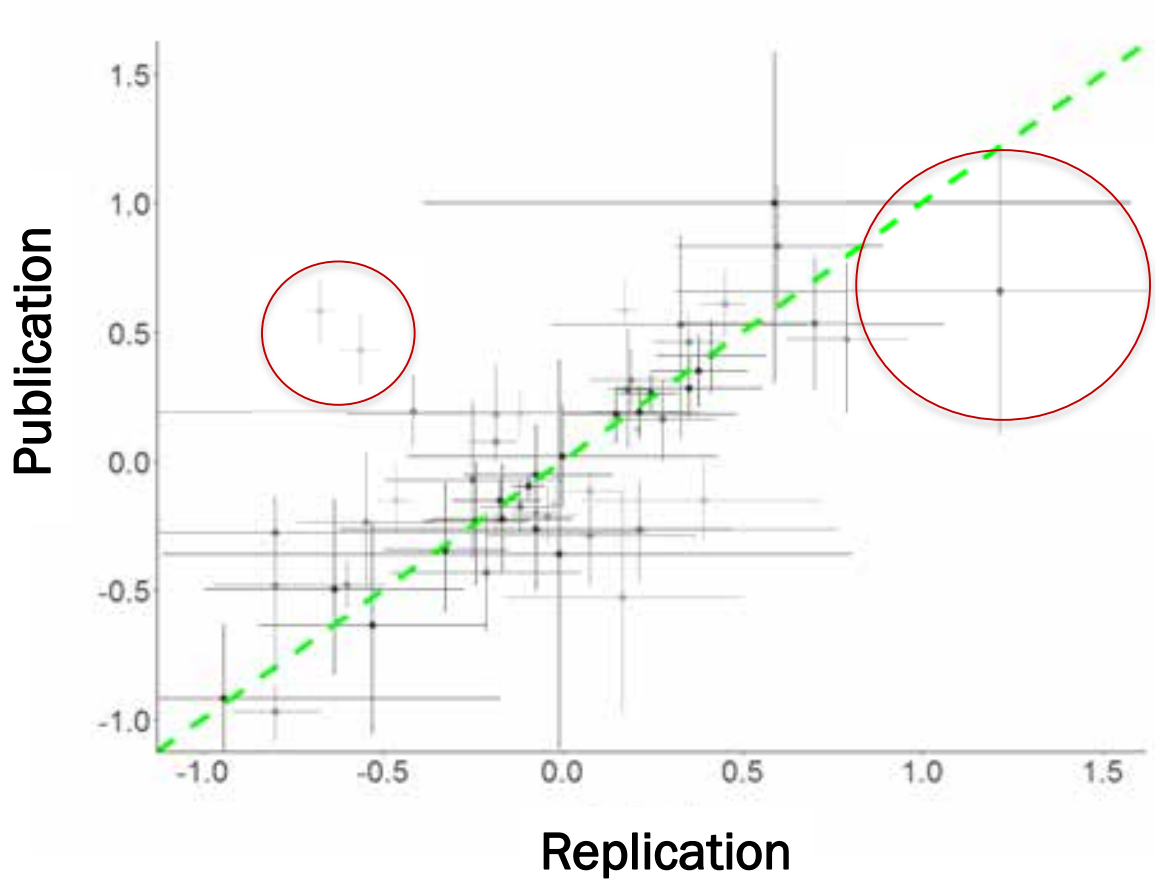


Effect Size and Confidence Limits

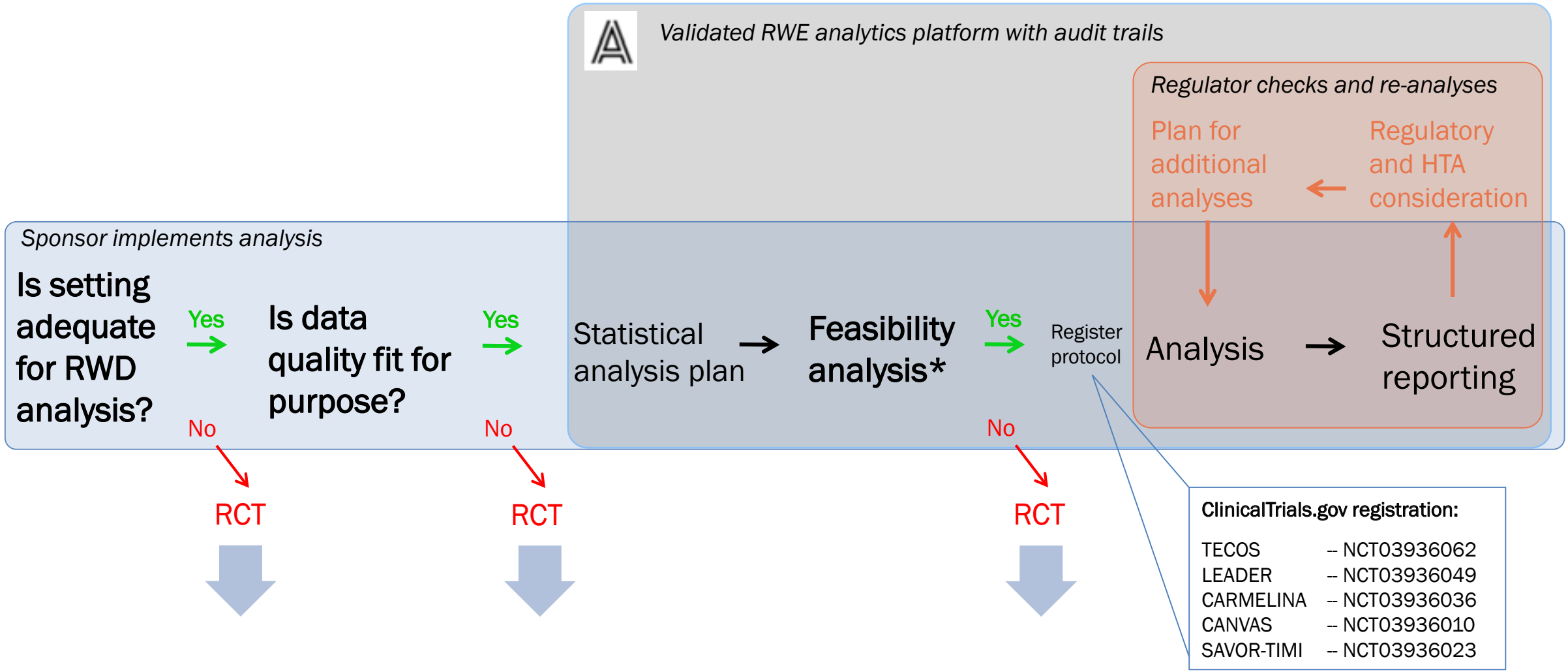
PI: Wang, Schneeweiss

Correlation between effect sizes
(publication vs replication)

- Correlation coefficient = 0.74



A pathway with regulatory validation



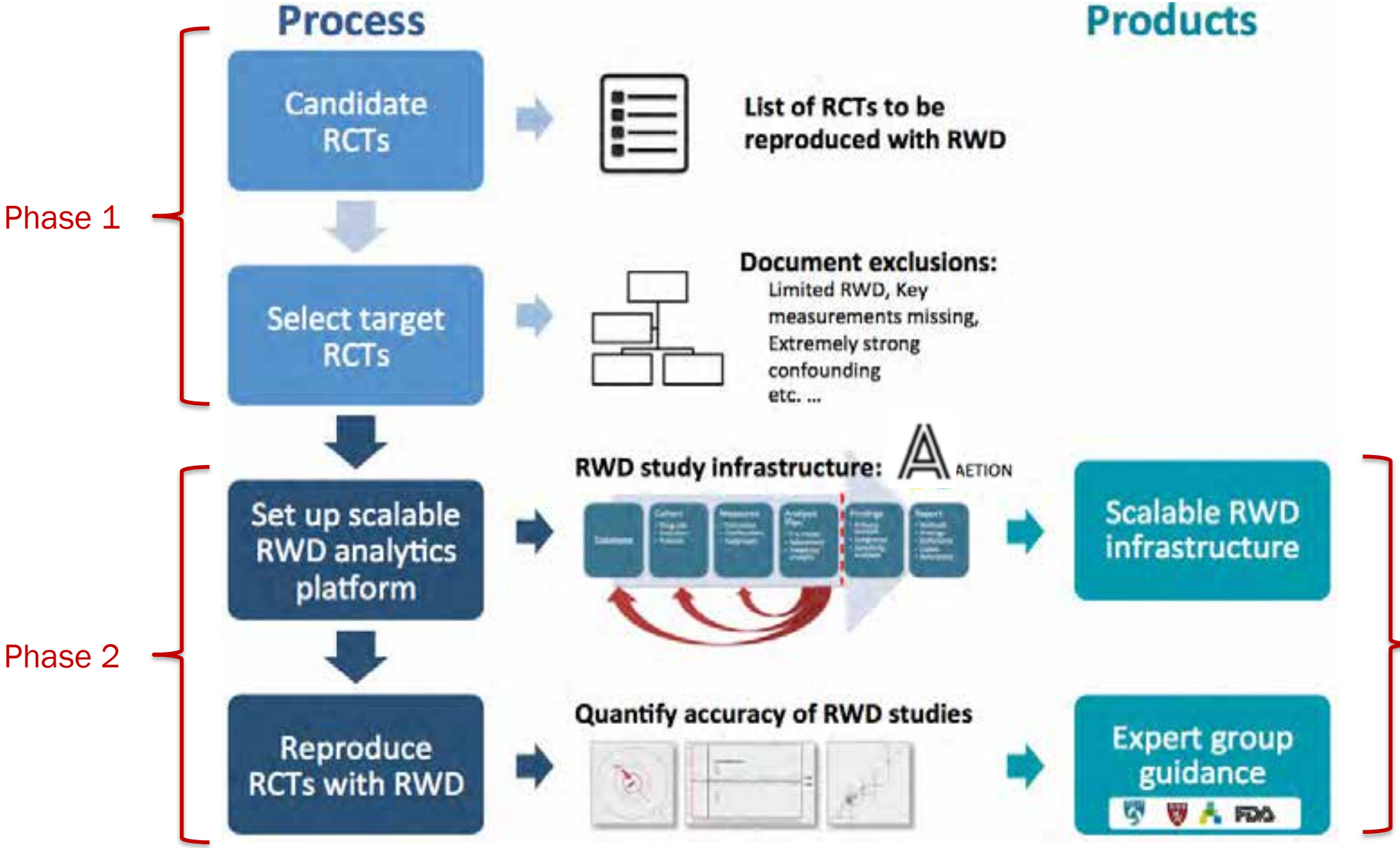
Franklin, Glynn, Martin, Schneeweiss. CPT 2019

* Feasibility analysis can include 1) checking covariate balance after applying the chosen confounding adjustment strategy, 2) checking statistical power, 3) evaluating positive or negative control outcomes, and 4) other analyses, *without* evaluating the study outcomes in the two treatment groups.

How well can RWD analyses reproduce RCT findings?

PI: Franklin, Schneeweiss

30 +7



Franklin, Pawar, Martin, Glynn, Levenson, Temple, Schneeweiss. CP&T 2019

PI: Franklin, Schneeweiss

Database Study

followed by →

RCT

Using Real-World Data to Predict Findings of an Ongoing Phase IV Cardiovascular Outcome Trial: Cardiovascular Safety of Linagliptin Versus Glimepiride

Elisabetta Patorno,[1] Sebastian Schneeweiss,[1] Chandrasekar Gopalakrishnan,[1] David Martin,[2] and Jessica M. Franklin[1]

Diabetes Care 2019;42:1–7 | https://doi.org/10.2337/dc19-0069

Risk of CV events (3P-MACE)

HR = 0.91 (0.79 – 1.05)

NCT03648424

Design and baseline characteristics of the CARdiovascular Outcome Trial of LINAgliptin Versus Glimepiride in Type 2 Diabetes (CAROLINA®)

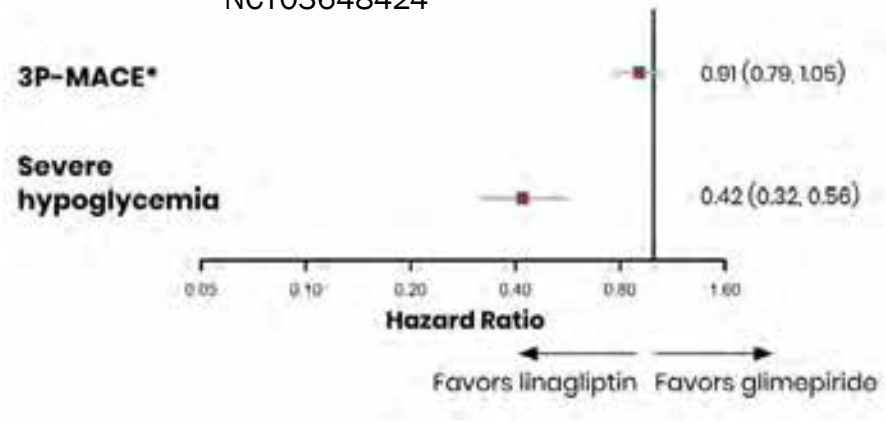
Nikolaus Marx[1], Julio Rosenstock[2], Steven E Kahn[3], Bernard Zinman[4,5], John J Kastelein[6], John M Lachin[7], Mark A Espeland[8], Erich Bluhmki[9], Michaela Mattheus[10], Bart Ryckaert[11], Sanjay Patel[12], Odd Erik Johansen[13] and Hans-Juergen Woerle[10]

CAROLINA

Risk of CV events (3P-MACE)

HR = ???

NCT01243424



Patorno E. et al. Diab Care 2019;42: June 14 epub

PI: Franklin, Schneeweiss

Database Study

followed by →

RCT

Using Real-World Data to Predict Findings of an Ongoing Phase IV Cardiovascular Outcome Trial: Cardiovascular Safety of Linagliptin Versus Glimepiride

Elisabetta Patorno,[1] Sebastian Schneeweiss,[1] Chandrasekar Gopalakrishnan,[1] David Martin,[2] and Jessica M. Franklin[1]

Diabetes Care 2019;42:1–7 | https://doi.org/10.2337/dc19-0069

Design and baseline characteristics of the CARdiovascular Outcome Trial of LINAgliptin Versus Glimepiride in Type 2 Diabetes (CAROLINA®)

ADA June 10, 2019

Nikolaus Marx[1], Julio Rosenstock[2], Steven E Kahn[3], Bernard Zinman[4,5], John J Kastelein[6], John M Lachin[7], Mark A Espeland[8], Erich Bluhmki[9], Michaela Mattheus[10], Bart Ryckaert[11], Sanjay Patel[12], Odd Erik Johansen[13] and Hans-Juergen Woerle[10]

CAROLINA

Risk of CV events (3P-MACE)

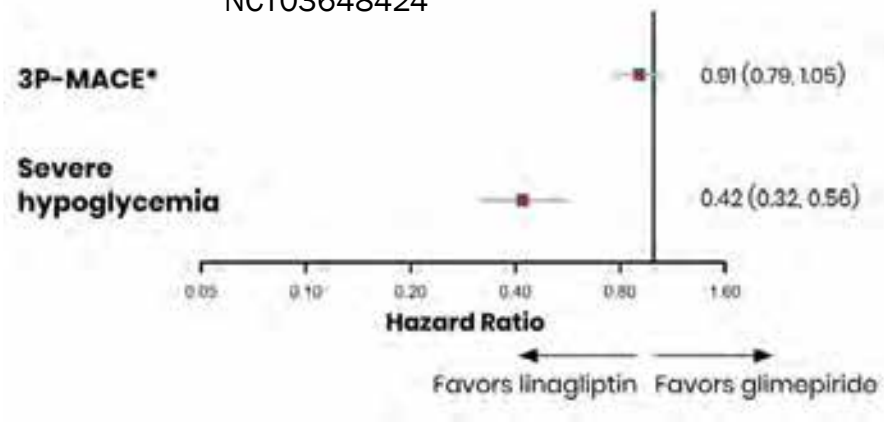
HR = 0.91 (0.79 – 1.05)

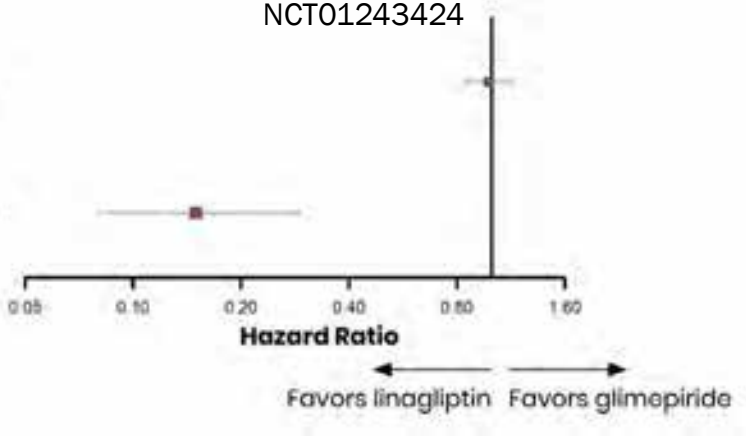
NCT03648424

Risk of CV events (3P-MACE)

HR = 0.98 (0.84 – 1.14)

NCT01243424

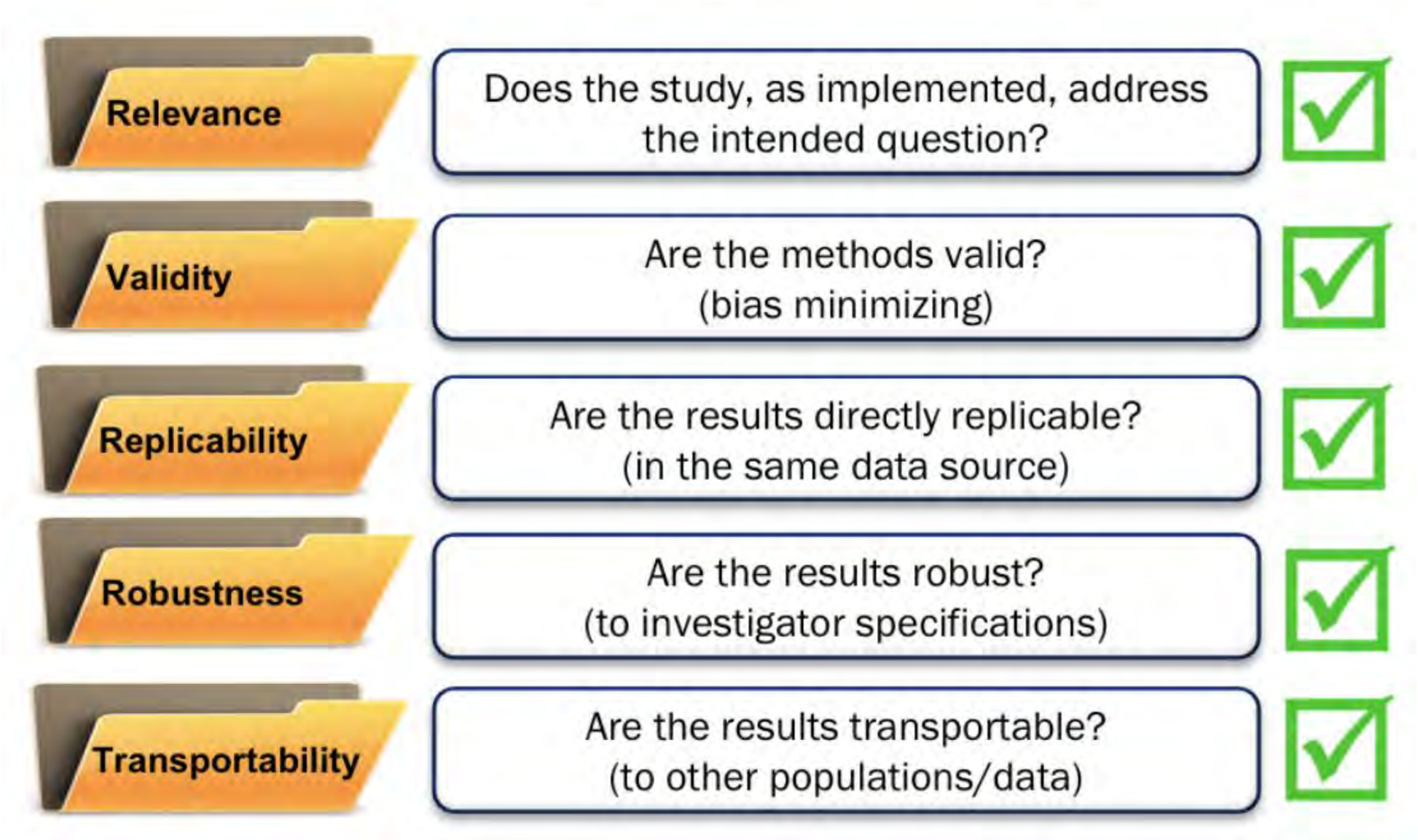




AETION

Patorno E. et al. Diab Care 2019;42: June 14 epub

Independent Evidence Dossiers for decision makers?



Session IV: Methodological and Analytical Considerations for Observational Studies

Substantial Evidence of Effectiveness Consists of Adequate and Well-Controlled Clinical Investigations

	Summary of Essential Characteristics of Adequate and Well-Controlled Investigations - 21 CFR 314.126
1	There is a clear statement of the objectives of the investigation and a summary of the proposed or actual methods of analysis in the protocol for the study and in the report of its results.
2	The study uses a design that permits valid comparison w/ a control to provide a quantitative assessment of drug effect.
3	The method of selection of subjects provides adequate assurance that they have the disease or condition being studied, or evidence of susceptibility and exposure to the condition against which prophylaxis is directed.
4	**The method of assigning patients to treatment and control groups minimizes bias and is intended to assure comparability of the groups with respect to pertinent variables, such as age, sex, severity of disease, duration of disease, and use of drugs or therapy other than the test drug. Ordinarily…assignment is by randomization.**
5	Adequate measures are taken to minimize bias on the part of the subjects, observers, and analysts of the data.
6	The methods of assessment of subjects' response are well-defined and reliable.
7	**There is an analysis of the results of the study adequate to assess the effects of the drug. The report of the study should describe the results and the analytic methods used to evaluating them, including any appropriate statistical methods. The analysis should assess…the comparability of test and control groups with respective to pertinent variables.**

Analytical Methods for Addressing Unmeasured Confounding in Observational Studies of Treatment Effectiveness

- Standard bias control methods assume “no unmeasured confounding”
- Will unmeasured confounder(s) be strong enough to create bias based on quantitative assessment?
- If yes, then conduct sensitivity analysis to evaluate the impact of unmeasured confounding

No **information on unmeasured confounder(s)**	***Internal*** **information on unmeasured confounder(s)**	***External*** **information on unmeasured confounder(s)**
Plausibility assessment set I: Negative control Pseudo treatment Manski’s partial identification Empirical distribution calibration	**Plausibility assessment set II:** Plausibility assessment set I + Rosenbaum-Rubin sensitivity analysis Rosenbaum sensitivity analysis	**Plausibility assessment set III:** Plausibility assessment set I + Rosenbaum-Rubin sensitivity analysis Rosenbaum sensitivity analysis
Adjusted analysis set I: Instrumental variable Regression discontinuity Difference in difference method Missing cause approach Trend-in-trend method Perturbation variable	**Adjusted analysis set II:** Adjusted analysis set I + Bayesian twin regression Multiple imputation Propensity score calibration	**Adjusted analysis set III:** Adjusted analysis set I + Bayesian twin regression Propensity score calibration

Zhang X, Faries DE, Li H, Stamey JD, Imbens GW. Pharmacoepidemiology and Drug Safety, 2018; 27(4):373-382,

Session IV: Methodological and Analytical Considerations for Observational Studies

Break

Session V: Opportunities to Ascertain Endpoints in Routine Clinical Care Settings

rwEndpoints Use Case: Assessing Frontline Treatment Regimens in Real-world Patients with Advanced Non-Small Cell Lung Cancer

Jeff Allen, PhD

Friends of Cancer Research

Background and Pilot 1.0 Findings

Initial Pilot Project Focus:

- Evaluated the performance of real-world endpoints across multiple data sets by focusing on a common question: **What outcomes can be evaluated for advanced NSCLC (aNSCLC) patients treated with immune checkpoint inhibitors?**

- **Can real-world endpoints be used to accurately characterize differences between available interventions?**
- **Can further alignment on data quality and standards be used to develop an analytic framework to evaluate real world endpoints?**

treated with immune checkpoint inhibitors from diverse RWD sources.

- Several extractable endpoints from EHR and claims data correlate with OS. Survival among patients as assessed through EHR and claims data fall within the range of median OS values observed in several immune checkpoint inhibitor trials.

Pilot 2.0: Establishing a Framework to Evaluate Real-World Endpoints

Project Goals: Explore potential endpoints that may be fit for assessing long term benefits of a product compared to an existing alternative	
Project Focus	What is the ability of different real-world endpoints (rwOS, rwTTD, rwTTNT, and rwPFS) to reflect effectiveness previously observed in clinical trials across two frontline treatment pairs in advanced non-small cell lung cancer (aNSCLC) patients?
Research Objectives	**Objective 1:** Description of demographic and clinical characteristics of patients with aNSCLC receiving Frontline doublet chemotherapy, PD-(L)1 monotherapy; or PD-(L)1 + doublet chemotherapy. **Objective 2:** Evaluate and compare rwOS, rwTTD, rwTTNT, and rwPFS among select frontline therapy pairs in aNSCLC patients: • Doublet chemotherapy versus PD-(L)1 monotherapy • Doublet chemotherapy versus PD-(L)1 + doublet chemotherapy
Study Design	This is a retrospective observational analysis of data derived from electronic health record (EHR) and claims based databases. The datasets generated for the study will include all relevant, retrospective patient-level HIPAA-compliant de-identified data available for eligible individuals up to a single specific data cutoff date of March 31, 2018.
Data Partners	**ASCO CancerLinQ/Concerto HealthAI, COTA, Flatiron Health, IQVIA, Kaiser Permanente/CRN, Mayo Clinic/OptumLabs®, McKesson, SEER, Syapse, and Tempus**

Real-World Endpoint Assessment

Real-world derived endpoint definitions

Overall survival (OS)

- *Data definition / computation*: Length of time from the index date to the date of death, or disenrollment, or last structured recorded clinical activity within the network or prescription, office or institutional billing claims data, or end of follow-up period, whichever occurs earliest. For claims data, health plan disenrollment date are incorporated if deaths are not captured among those who leave health plan coverage.

Time to Next Treatment (TTNT)

- *Data definition / computation*: Length of time from the index date to the date the patient received an administration of their next systemic treatment regimen or to their date of death if there is a death prior to having another systemic treatment regimen.

Time to Treatment Discontinuation (TTD)

- *Data definition / computation*: Length of time from the index date to the date the patient discontinues frontline treatment. The frontline treatment discontinuation date is defined as the last administration or non-cancelled order of a drug contained within the same frontline regimen.

Definition of progression in aNSCLC as evident in the EHR

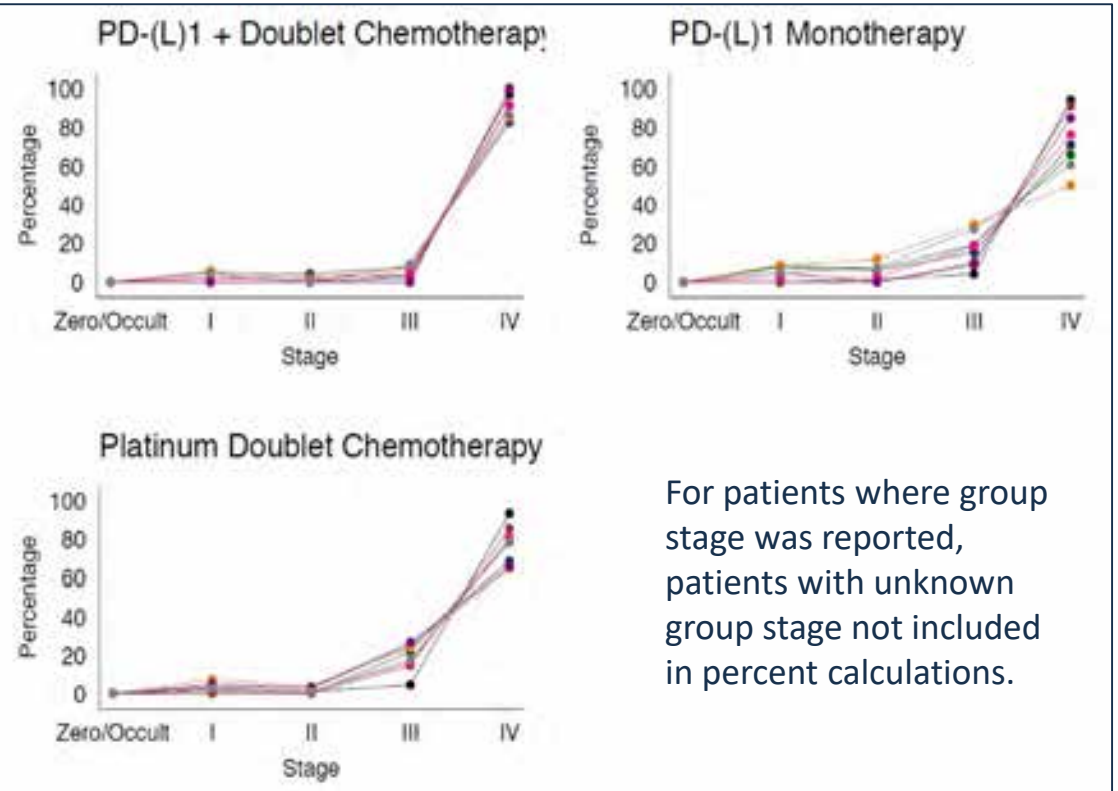
A **progression event** is a distinct episode in which the treating clinician concludes that there has been growth or worsening in the aNSCLC. The progression event (and date) is based on review of the patient chart.

Progression Free Survival (PFS)

- *Data definition / computation*: Length of time from the date the patient initiates frontline treatment (from the date the patient received administration for the first product in their frontline treatment) to the date of a rwP event, at least 14 days after frontline treatment initiation, or death. For patients without a rwP event prior to TTNT, rwPFS will be censored at the date of rwTTNT. For patients without a rwP event or a rwTTNT event and at least 180 days follow-up from last frontline treatment, rwPFS will be censored at rwTTD. Patients with a rwP event within 14 days from frontline treatment initiation will be excluded.

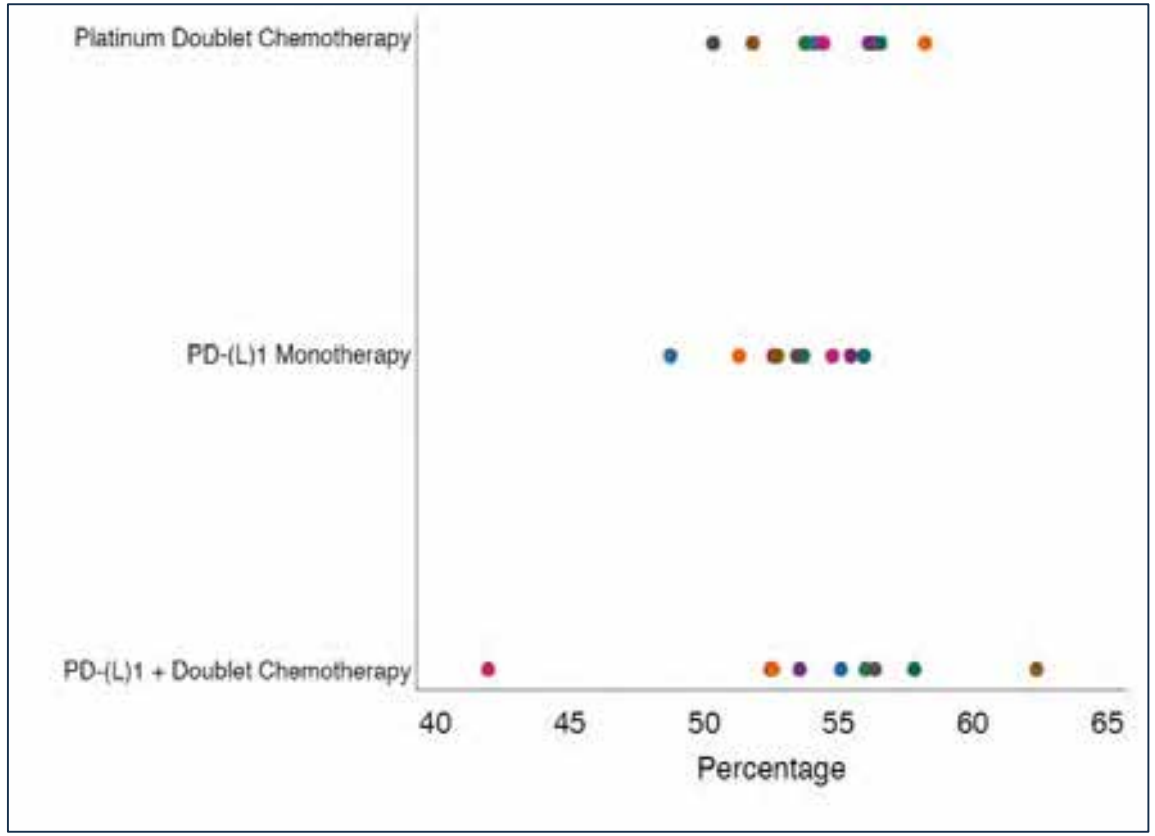
Demographic and Clinical Characteristics

Group Stage of Patients with aNSCLC Per Treatment Category



For patients where group stage was reported, patients with unknown group stage not included in percent calculations.

Percentage of Male aNSCLC Patients

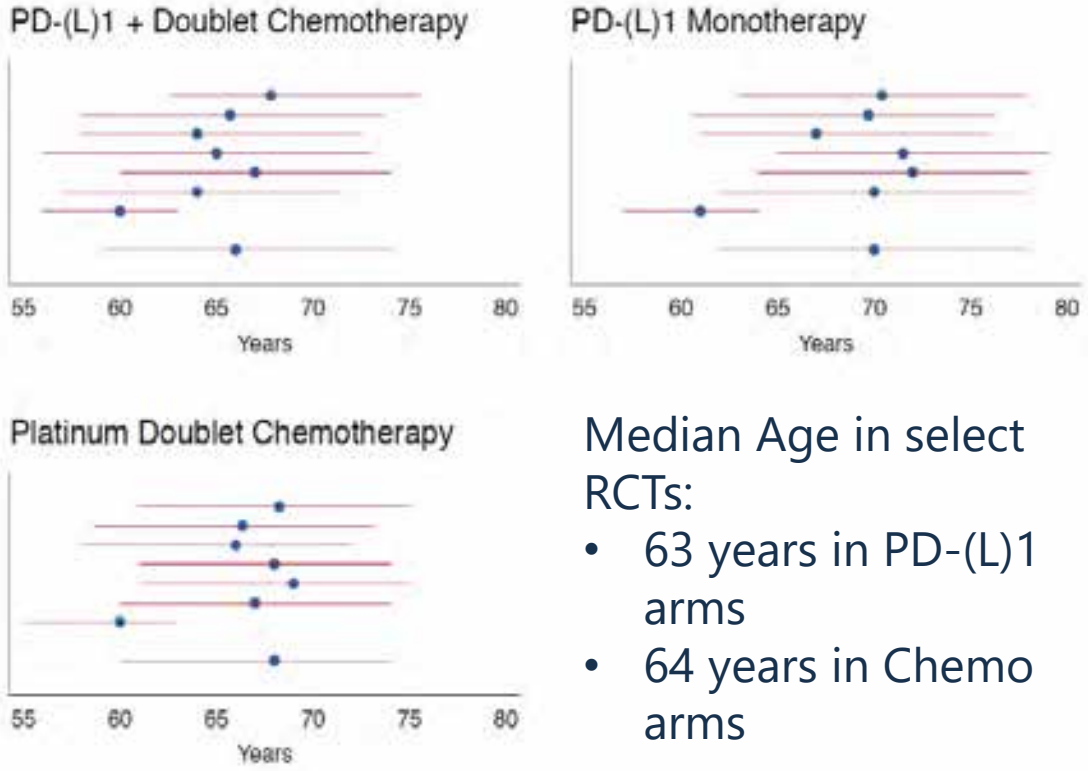


- Graphs are based on structured or unstructured information depending on the data source
- Graphs represent data of patients with values reported. Missing/unknown data are not represented in these graphs

FRIENDS of CANCER RESEARCH

Demographic and Clinical Characteristics

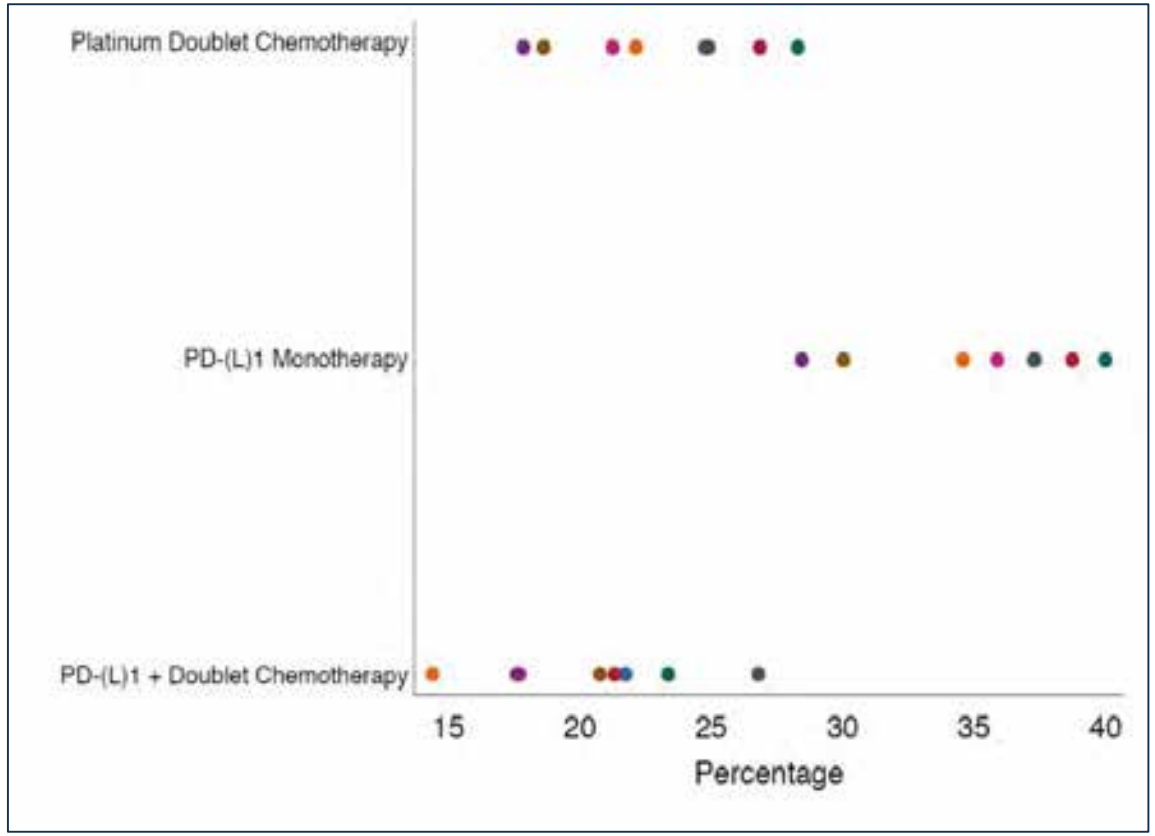
Median and Lower/Upper Quartiles of Age at Index



Median Age in select RCTs:

- 63 years in PD-(L)1 arms
- 64 years in Chemo arms

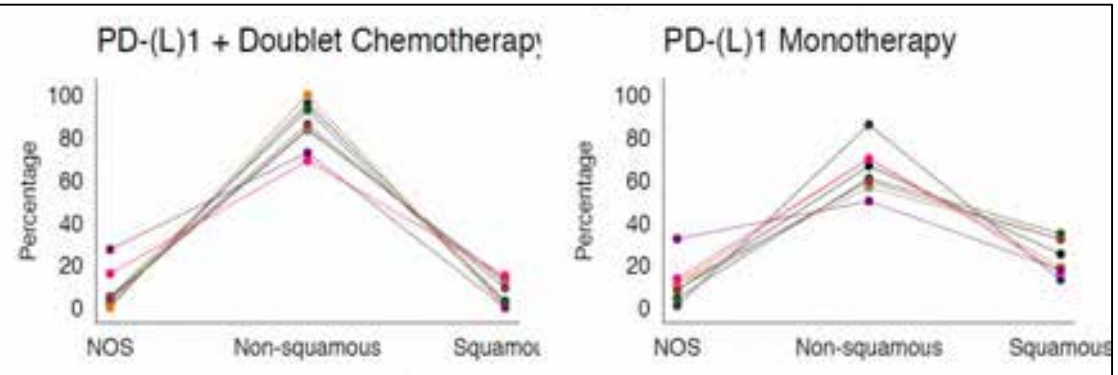
Percentage of aNSCLC Patients Age 75 or Older at Index



FRIENDS of CANCER RESEARCH

Demographic and Clinical Characteristics

Histology of Patients with aNSCLC Per Treatment Category

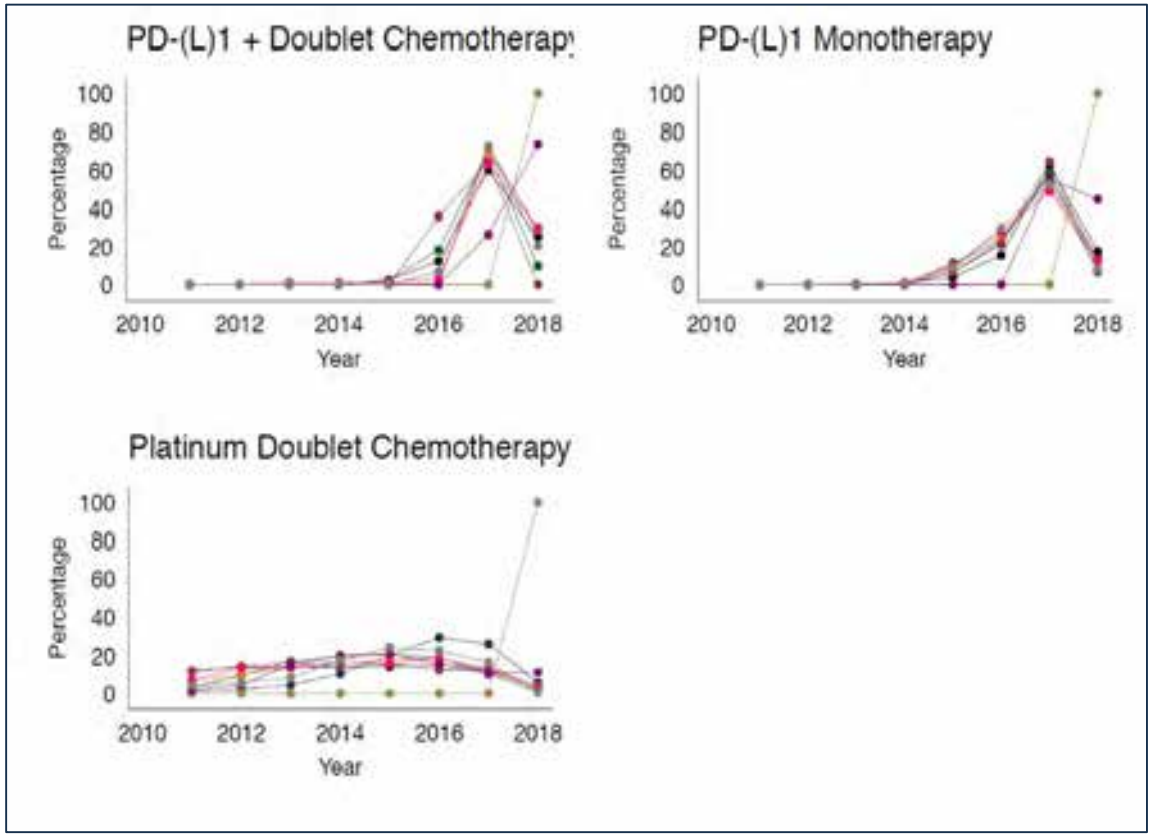


Histology of patients from select RCTs:

- Nonsquamous: 81.2%
- Squamous: 18.8%

NOS: Histology not otherwise specified

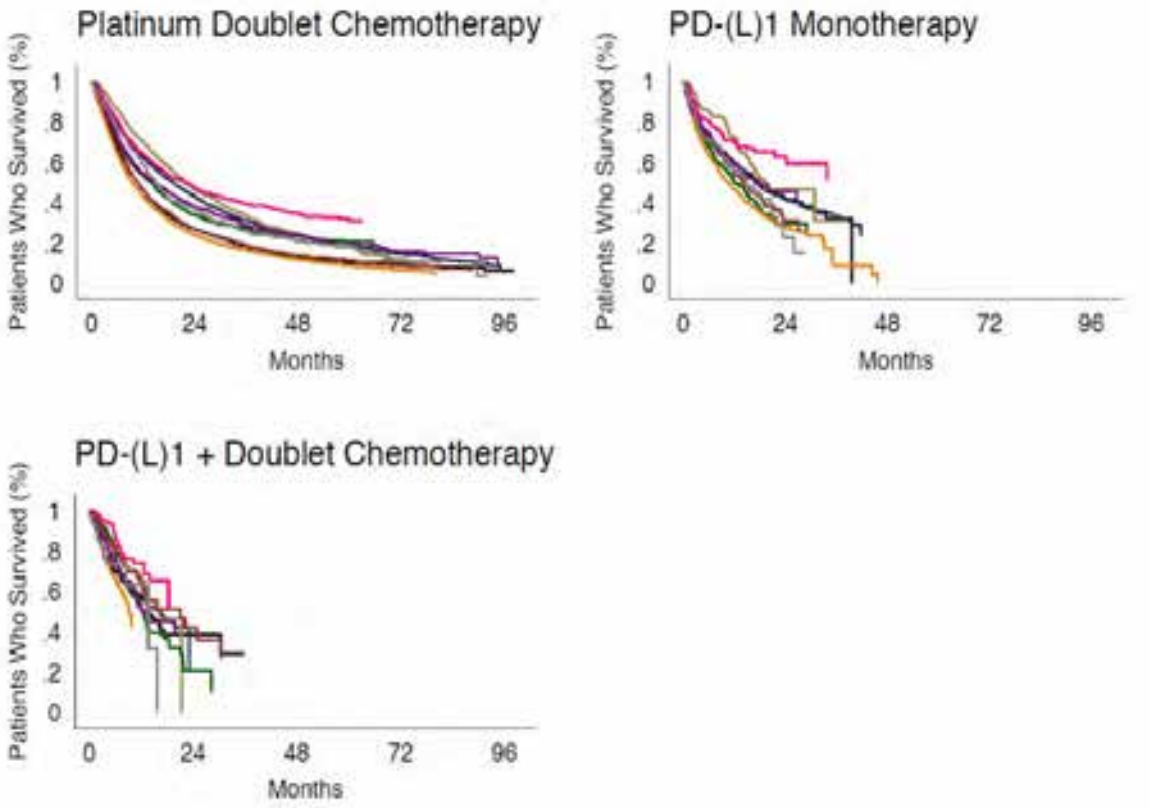
Year of Index Date Per Treatment Category



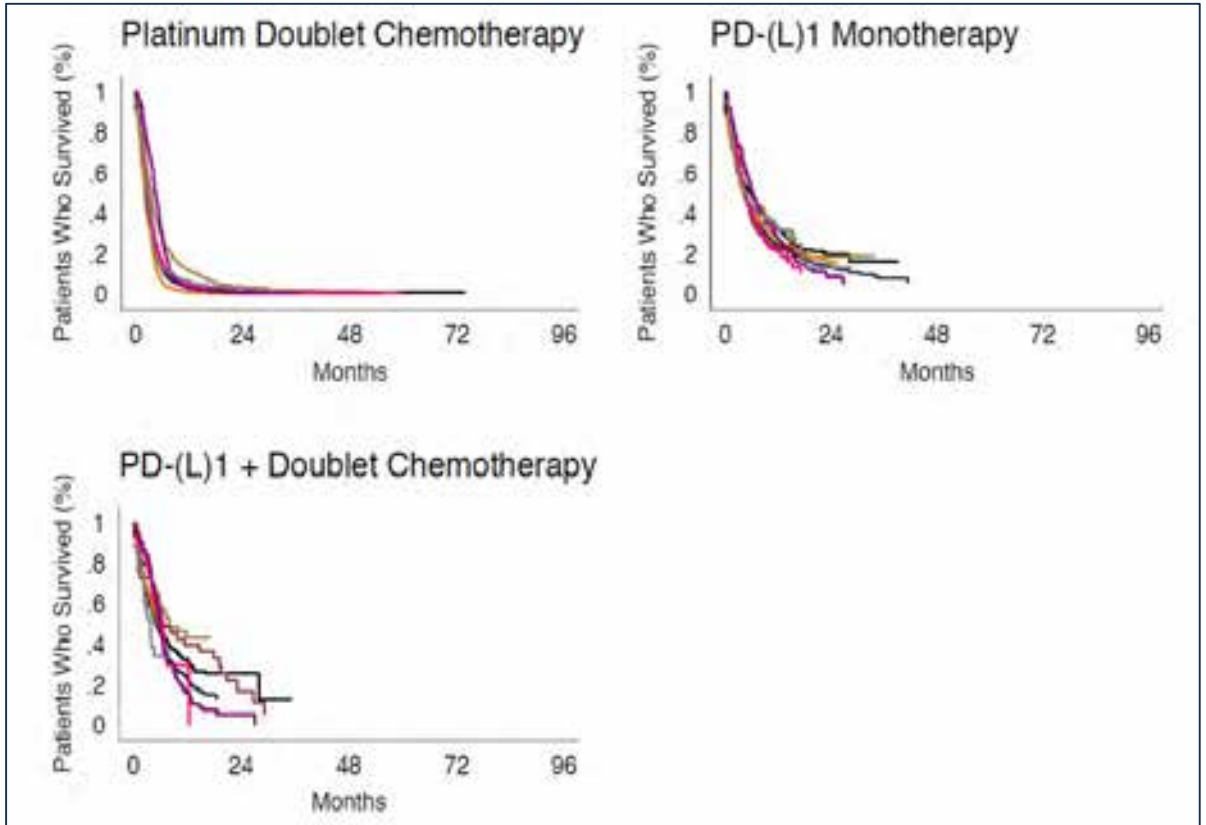
FRIENDS of CANCER RESEARCH

Kaplan Meier Curves per Treatment Group

rwOS

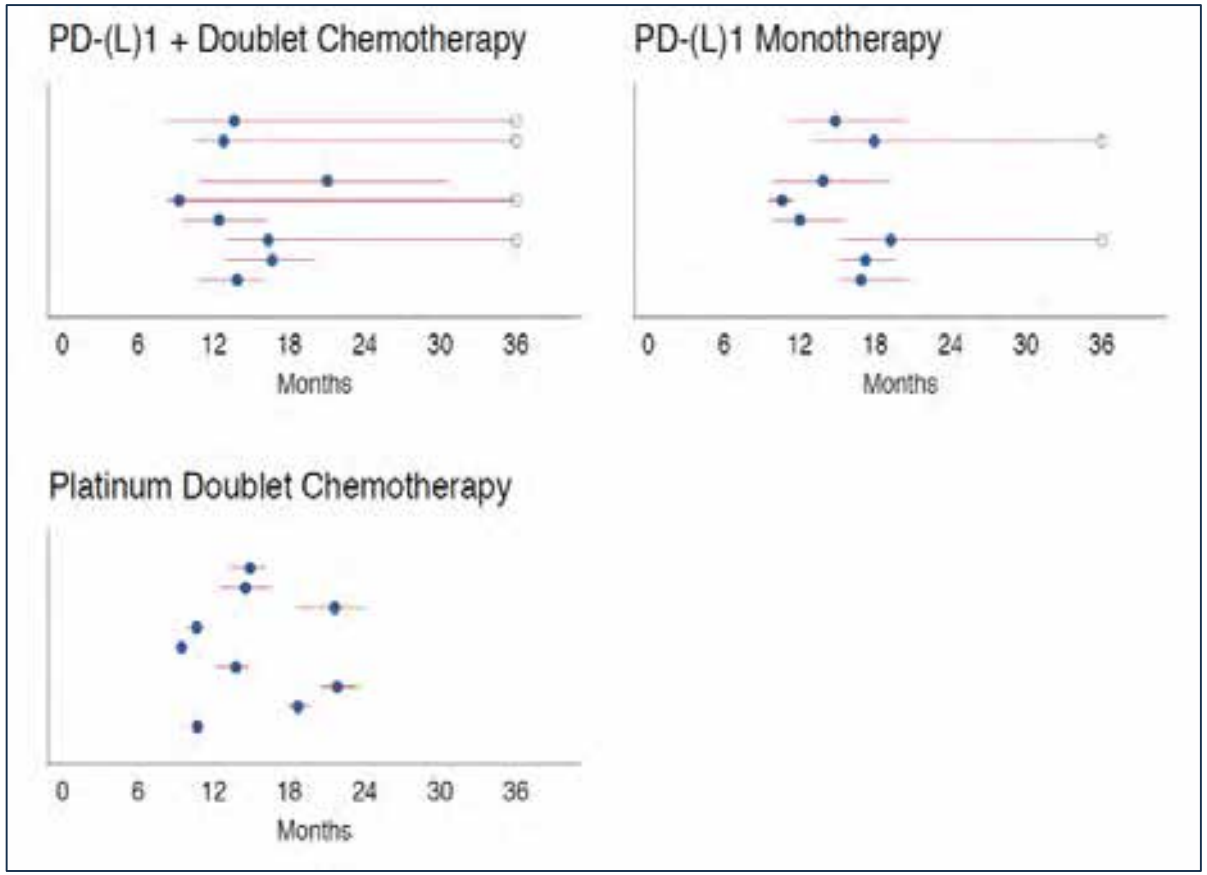


rwTTD

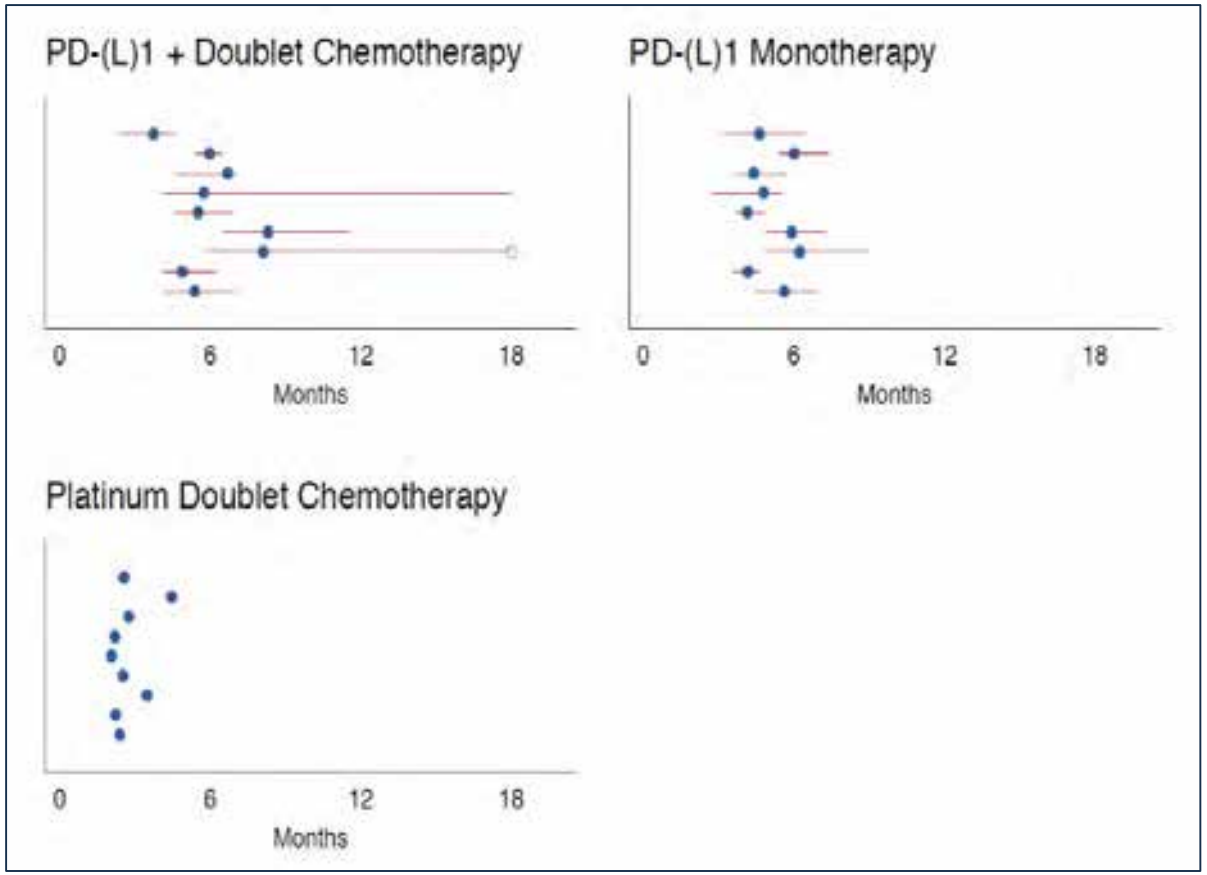


Estimates of Median Time per Treatment Category

rwOS



rwTTD



Conclusions

- It is possible to coordinate the efforts across numerous real-world oncology data organizations to reach high-level alignment on important data elements and definitions for real-world endpoints in the context of a focused research question.

- The depth of data varied across data providers and distinct characteristics were identified among the cohorts provided by each organization, likely attributable to the characteristics of the data source and the underlying population it is capturing.

- The results of this phase of the pilot project highlighted the ability to show differences in important prognostic demographic as well as clinical characteristics between trial patients and heterogenous real-world patient populations (e.g., median age, histology).

- It also demonstrated the ability to provide insight into recent trends in clinical care.

Next Steps

- Carefully review data and assess potential differences in the population characteristics, data source, and/or subtle differences in methodological assumptions made during the analysis that could impact outcomes

- Evaluate Treatment effect size in frontline therapy regimens using real-world endpoints
  - Stratified analyses: PD-(L)1 status, other patient demographics

- Conduct analysis among real-world patients that match RCT eligibility requirements in order to assess comparability to clinical trial populations. Such analyses may:
  - Help identify sources of variability – data source, treatment settings, provider level variation
  - Model methodology for potential data quality control
  - Inform a framework to assess performance of real-world endpoints

Acknowledgements

Pilot 2.0 Data Partners

- ASCO CancerLinQ/Concerto HealthAI
- COTA
- Flatiron Health
- IQVIA
- Kaiser Permanente/Cancer Research Network
- Mayo Clinic/OptumLabs®
- McKesson
- SEER-Medicare Linked Database
- Syapse
- Tempus

Key Collaborators

- FDA
- NCI

Data Analytics and Graph Generation

- Aetion

Session V: Opportunities to Ascertain Endpoints in Routine Clinical Care Settings

Open Comment Period

Closing Remarks

Adjournment