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

National Press Club
529 14th St NW, Washington, DC 20045
October 3, 2019

Join the conversation with #RWE2019
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%

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.
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

Diagnoses

- Demographics
- Medication orders
- Labs

Procedures

- Death data

Labs

- Geocodes
- Tumor registry
- Natural language processing-derived concepts
- Biosamples
- Claims

Patient-reported outcomes

Patient-generated data

Genomic results

Social determinants of health
#2 People-Centeredness

**Direct to Consumer**
- Flexible
- Frictionless
- Fun

**Direct to Participant**
- Personalized
- Streamlined
- Valuable
User-Reported Data
What people say

Task-Based Measures
Measures effort and physiology

Passive Sensing
What people actually do day to day
A Real World Example
Looking Back at a Disruptive Technology

“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
ADAPTABLE: What’s the Right Dose of Aspirin?

eScreening, eEnrollment and eFollow-up

Call FOLLOW-UP
- Patient Reported Outcomes
- Medication use
- Health outcomes

Portal FOLLOW-UP
- Patient Reported Outcomes
- Medication use
- Health outcomes

PCORNet Coordinating Center FOLLOW-UP
- Via Common Data Model
- Longitudinal health outcomes

CMS, Payer, FOLLOW-UP
- Longitudinal health outcomes

N= 15,000

http://adaptablepatient.com
There are 5 steps to join the study!

Watch
the ADAPTABLE short video

Read
more details about participating in ADAPTABLE

Answer
a few questions about the study

Join
the ADAPTABLE study

Inform
us about your current health

LET'S GET STARTED!
<table>
<thead>
<tr>
<th>Pre-study</th>
<th>Study setup</th>
<th>Recruitment</th>
<th>Study conduct</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Assess sites’ use of EHR to facilitate research</td>
<td>• Utilize EHR to identify local participants</td>
<td>• E-consent with comprehension questions</td>
<td>• Trials specific data capture from care delivery</td>
</tr>
<tr>
<td>• Usability of inclusion and exclusion criteria</td>
<td>• Embed encounter instructions and site content into EHR</td>
<td>• Incorporate screening criteria into EHR for</td>
<td>• Auto-populated CRFs fields from EHR</td>
</tr>
<tr>
<td>• Refine protocol</td>
<td>• Pre-consent &amp; study specific consent</td>
<td>- Scheduling patients</td>
<td>• Extract data to facilitate work of study coordinator</td>
</tr>
<tr>
<td>• Community interaction profiles with health system</td>
<td>• Alert clinician about trial</td>
<td>- Contacting patients</td>
<td>• Query data to identify events</td>
</tr>
<tr>
<td>• Feasibility analysis</td>
<td>• Model outcomes</td>
<td>- Recruiting patients</td>
<td>• Participant retention and education</td>
</tr>
<tr>
<td>• Recruitment plan</td>
<td></td>
<td>• Alert clinician of patient eligibility</td>
<td>• Return of results</td>
</tr>
</tbody>
</table>

- EHR Health Portals
  - Patient opt in/out for types of studies
Important Matters
Quality & Outcomes
What’s the Purpose?

- New Drug Approval
- Label Expansions and Revisions
- Post Market Commitments
- Clinical Guidance
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
- $$\text{is limitless}$$

**Real World**
- Routine Population
- Usual Care
- Unblinded
- Active control
- Passive data collection
- Participant directed data collection
- $$\text{leveraged with embedded trials}$$

**Hybrid**
The Puzzle Coming Together?

Key areas, preferences & questions

Electronic Health Data

Health Systems & Communities

Engagement | Enrollment

EHR extraction

PRO

Common Data Model

Study data

Analysis/Results

Evaluation & Feedback/Dissemination

Notifications and Messaging

Devices, surveys, wearables, etc.

Personalized Health Initiatives

Return of Results

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)?

• 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)?
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?

Issues raised:

- 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

<table>
<thead>
<tr>
<th>Randomized Interventional</th>
<th>Interventional non-randomized</th>
<th>Non-randomized / non-interventional</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Traditional Randomized Trial Using RWD Elements</strong></td>
<td><strong>Trials in Clinical Practice Settings</strong></td>
<td><strong>Observational Studies</strong></td>
</tr>
<tr>
<td>RWE to assess enrollment criteria / trial feasibility</td>
<td>eCRF + selected outcomes identified using EHR/claims data</td>
<td>Prospective data collection</td>
</tr>
<tr>
<td>RWE to support site selection</td>
<td>Mobile technology used to capture supportive endpoints (e.g., to assess ambulation)</td>
<td>Registry trials/study</td>
</tr>
</tbody>
</table>

- **Pragmatic RCTs**
  - Pragmatic RCT using eCRF (+/- EHR data)
  - Pragmatic RCT using claims and EHR data

- **A large, simple trial**
- **A pragmatic trial**

- **Single arm study using external control**

**Increasing reliance on RWD**

- **Traditional RCT**
- **RWE / pragmatic RCTs**
- **Observational cohort**
Pragmatic randomized clinical trials: an overview of components

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

**Recruitment:** patients in practice settings

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

**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

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

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

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

**Primary analysis:** usually inclusive

**Do the patients actually have the targeted disease?**

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

**How well are we detecting efficacy endpoints and safety?**

**How well are we detecting the effect if patients do not take the drug? Is adherence per se an issue?**

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

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

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
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

Data Entry

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

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

Continuum of breast patient care

Screening

Other Diagnostics

Treatment Planning

Treatment

Follow-up / Survivorship

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.

Research Coordinator entry
- CRFs

TRIALS / REGISTRIES

• 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

PATIENT-REPORTED DATA
- Health History
- Social History
- Health Habits
- Family History
- Symptoms
- Quality of life

CLINICAL
- Diagnostic Findings
  - Clinical Exam
  - Imaging
  - Biopsy Pathology
  - Staging
- Clinical Trial Matching
- Treatment
  - Surgery
  - Systemic
  - Radiation
- Adverse Events

DATA ACCESS
- Single-source of truth
- Consistency among secondary users

FINANCE/BILLING
- Billing Pre-Authorization
- Billing (Institution)
- Payers (Insurance)

RESEARCH
- Less cleaning required
- Reduced Staffing

ADMINISTRATION
- ACS Quality Auditor
- Admin Analyst
- Inst. Quality Auditor
- Value Chain Analyst

Point of Care Data Collection
Process reengineering
Starting with the AS IS and working towards the TO BE
### “Enter Once, Use Many”

<table>
<thead>
<tr>
<th>Domain / Data Elements</th>
<th>New PATIENT (AT ANY PHASE)</th>
<th>SCREENING</th>
<th>DIAGNOSTIC</th>
<th>TREATMENT PLANNING</th>
<th>SURGERY</th>
<th>SYSTEMATIC TREATMENT</th>
<th>RADIATION TREATMENT</th>
<th>FOLLOW-UP</th>
<th>RESEARCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient-Reported Outcomes</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
</tr>
<tr>
<td>Patient Health History</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
</tr>
<tr>
<td>Imaging</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
</tr>
<tr>
<td>Biopsy Pathology</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
</tr>
<tr>
<td>Clinical Exam and Stage</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
</tr>
<tr>
<td>Clinical Trial Matching</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
</tr>
<tr>
<td>Treatment</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
</tr>
<tr>
<td>Final Pathology</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
</tr>
</tbody>
</table>

- **New data**  
- **Confirm or Update**  
- **View Only**  
- **Confirm/Additional Data Added**
Supporting clinical trials and data submissions

Electronic Health Record Systems
Decision Support at Point of Care

Clinical Trial Data Submissions & Standards
Mobile device - Patient Reported Outcomes, Adverse Event Reporting

- TAUG-BrCa
- S(AE)
- CTCAE
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
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

PRACTICE TYPE DISTRIBUTION (SIGNED)

15% Academic

26% Hospital/Health system

59% Private/Independent
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/
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
Session I: Establishing a High-Quality RWD Ecosystem

Join the conversation with #RWE2019
Session I: Establishing a High-Quality RWD Ecosystem

Nancy Yu
CEO, RDMD

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

<table>
<thead>
<tr>
<th>Challenge</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard of care is often poorly defined or not broadly adopted</td>
<td>Variability in analyzing data across different sites</td>
</tr>
<tr>
<td>Clinical outcomes assessments may 1) not be used, 2) anecdotally used, or 3) inconsistently / subjectively recorded</td>
<td>Incomplete data used to inform endpoint validation</td>
</tr>
<tr>
<td>Limited overall understanding of conditions to interpret complex clinical data</td>
<td>Difficulty in developing standard policies &amp; procedures</td>
</tr>
<tr>
<td>Curated data is not equal to standardized data</td>
<td>Ensuring harmonization with existing standards is not always pragmatic</td>
</tr>
<tr>
<td>Dispersed populations requires data from disparate EHRs</td>
<td>Rigorous standards development &amp; quality control needed</td>
</tr>
</tbody>
</table>
Our technology platform enables end-to-end Data Quality Control & Data Relevancy.
Data Quality Control: Technology, processes, training

- **Patient E-consent**
  - E-signature capture
  - ID verification
  - Two-factor authentication
  - IRB update content management system

- **Record Requesting**
  - Hospital data verification
  - Automated API for e-fax & receipt
  - Digital & physical audit trails

- **Record Processing**
  - Source document review (ALCOA)
  - QC: document verification & classification
  - Records attributed to Participant & processing staff
  - System permissions based on user roles

- **Record Abstraction**
  - 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

- **Export & Analysis**
  - Data transfer / export checks
  - Periodic audits around consent & protocol scope
# Data Quality Control: Data abstraction conducted under a central research protocol

<table>
<thead>
<tr>
<th>Trained abstractors</th>
<th>RDMD technology platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trained abstractors with clinical research or nurse practitioner backgrounds</td>
<td>Software enables effective document review &amp; data capture in predetermined forms</td>
</tr>
</tbody>
</table>

**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
Pre-programmed forms

Instant structured output

Abstractor data capture

Evidence linking

Pre-programmed forms
**Data Relevancy:** Growing clinical module library maps to industry standards

<table>
<thead>
<tr>
<th>Standard Modules</th>
<th>Therapeutic-Area Specific Modules</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Examples:</strong></td>
<td></td>
</tr>
<tr>
<td>Diagnosis</td>
<td></td>
</tr>
<tr>
<td>Demographics</td>
<td>Clinical milestone modules</td>
</tr>
<tr>
<td>Lab values</td>
<td>Disease-specific symptom modules</td>
</tr>
<tr>
<td>Echocardiograms</td>
<td>Disease-specific assessments</td>
</tr>
<tr>
<td>Concomitant medications</td>
<td>• Urine GAG testing, MPS enzyme testing</td>
</tr>
<tr>
<td>Genetic testing</td>
<td></td>
</tr>
<tr>
<td>EKGs</td>
<td></td>
</tr>
<tr>
<td>Healthcare utilization</td>
<td></td>
</tr>
</tbody>
</table>

**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

<table>
<thead>
<tr>
<th>High Confidence</th>
<th>Medium Confidence</th>
<th>Low Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original source documents available</td>
<td>Physician confirms endpoints in note, but source documents unavailable</td>
<td>Endpoints briefly referenced in physician note; source documents unavailable</td>
</tr>
</tbody>
</table>

**Patient Case**

- **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.

**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**
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*

**Early Natural History**

- Birth
- Hiatal hernia diagnosis & repair (x2)
- Hearing loss
- Met early gross & fine motor skills
- Chronic GI issues

**Diagnosis**

- MPS suspected
- Genetics evaluation & phenotyping
- Elevated urine MPS
- Official Hunter Syndrome diagnosis

**Management**

- Treatment & management initiated
- Neurocognitive & behavioral concerns
- Myopia
- Abnormal echo (cardiology)

*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

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
Signs of MPS discovered incidentally

Age 4:
Present to GI clinic for chronic diarrhea; X-ray ordered
Analysis of real-world data, *in context*, can reveal the earliest signs of rare conditions.

**Pre-Diagnosis**

**Birth**
Normal newborn screen, ABR

<table>
<thead>
<tr>
<th>Age 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-17 months:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENT Notes: 44  Audiograms: 9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-2 years:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signs of MPS discovered incidentally</td>
</tr>
</tbody>
</table>

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

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

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

**Ent Documentation of PE Tubes**

*Provider: Dr. X (Otolaryngology)*

Reason for Appointment
1. Ear infection

History of Present Illness

*Postop: Post-op app 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.
Analysis of real-world data, *in context*, can reveal the earliest signs of rare conditions

![Radiology report](image)

**Pre-Diagnosis**

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
Signs of MPS discovered incidentally

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

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

Pre-Diagnosis

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:
Signs of MPS discovered incidentally.

Age 4:
Present to GI clinic for chronic diarrhea; X-ray ordered.
Analysis of real-world data, *in context*, can reveal the earliest signs of rare conditions

**Birth**
Normal newborn screen, ABR

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

**Age 4:**
Signs of MPS discovered incidentally

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

**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.
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:** Urine glycosaminoglycans ordered.
- **Age 4:** Hunter diagnosis confirmed.

**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.

**Genetics Clinic Note**

**Quantitative Urine MPS (External Lab X)**

Mucopolysaccharides (MPS) \( \leq 16.0 \) 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.

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.

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
- Myopia diagnosed; no retinopathy

**Provider:** Dr. X (Otolaryngology)
**Encounter Date:** 2/9/2017

**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

**Age 6:**
- Bicuspid aortic valve, thickened mitral valve noted on echocardiogram
- Ongoing neurocognitive concerns; autism diagnosis
Analysis of the post-diagnostic journey allows for tracking of long-term outcomes.

**Age 4:**
- Initiation of IV idursulfase
- Myopia diagnosed; no retinopathy
- Ongoing monitoring via ENT, ophthalmology, cardiology, GI, orthopedics, developmental pediatrics, genetics

**Age 6:**
- Bicuspid aortic valve, thickened mitral valve noted on echocardiogram
- Myopia diagnosed

**OPHTHALMOLOGY CLINIC NOTE**

Provider: Dr. O (Ophthalmology)

**Encounter Date:** 12/5/2016

**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
- Myopia diagnosed; no retinopathy
- Neurocognitive & behavioral concerns noted; ADHD diagnosis

Age 6:
- Ongoing neurocognitive concerns; autism diagnosis
- Bicuspid aortic valve, thickened mitral valve noted on echocardiogram

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

<table>
<thead>
<tr>
<th>Score</th>
<th>SS</th>
<th>%ile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Scale IQ</td>
<td>101</td>
<td>53</td>
</tr>
<tr>
<td>Verbal Comprehension Composite</td>
<td>99</td>
<td>47</td>
</tr>
<tr>
<td>Visual Spatial Composite</td>
<td>97</td>
<td>42</td>
</tr>
<tr>
<td>Fluid Reasoning Composite</td>
<td>100</td>
<td>50</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Score</th>
<th>SS</th>
<th>%ile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Scale IQ</td>
<td>80</td>
<td>9</td>
</tr>
<tr>
<td>Verbal Comprehension Composite</td>
<td>81</td>
<td>10</td>
</tr>
<tr>
<td>Visual Spatial Composite</td>
<td>78</td>
<td>7</td>
</tr>
<tr>
<td>Fluid Reasoning Composite</td>
<td>85</td>
<td>16</td>
</tr>
</tbody>
</table>
Evaluation of multiple source documents is required to understand the patient journey in rare disease.
Thank you!

Nancy Yu, CEO
Kristina Cotter, PhD, CGC, MS, Research Director

nancy@rdmd.com
kristina@rdmd.com
Session I: Establishing a High-Quality RWD Ecosystem
Collaborating Organizations

**Lead – HPHC Institute**

- Department of Population Medicine
- Harvard Medical School
- Harvard Pilgrim Health Care Institute

**Data & Scientific Partners**

- Optum
- HealthCore
- Anthem
- HCA Healthcare
- Massachusetts General Hospital
- Department of Population Health Sciences
- Duke University School of Medicine
- Vanderbilt School of Medicine

**Scientific Partners**

- Penn Medicine
- IQVIA
- Kaiser Permanente
- Aetna
- UNC Gillings School of Global Public Health
- T.H. Chan School of Public Health
- University of Florida College of Pharmacy
- UAB
- UIC
- College of Public Health
- College of Pharmacy

https://www.sentinelinitiative.org/collaborators
### Available Data Elements

#### Administrative Data

<table>
<thead>
<tr>
<th>Registry Data</th>
<th>Inpatient Data</th>
<th>Clinical Data</th>
<th>Mother-Infant Linkage Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Enrollment</strong></td>
<td><strong>Demographic</strong></td>
<td><strong>Dispensing</strong></td>
<td><strong>Diagnosis</strong></td>
</tr>
<tr>
<td>Patient ID</td>
<td>Patient ID</td>
<td>Patient ID</td>
<td>Patient ID</td>
</tr>
<tr>
<td>Enrollment Start &amp; End Dates</td>
<td>Dispensing Date</td>
<td>Service Date(s)</td>
<td>Service Date(s)</td>
</tr>
<tr>
<td>Drug Coverage</td>
<td>National Drug Code (NDC)</td>
<td>Encounter ID</td>
<td>Encounter ID</td>
</tr>
<tr>
<td>Medical Coverage</td>
<td>Days Supply</td>
<td>Encounter Type and Provider</td>
<td>Encounter Type and Provider</td>
</tr>
<tr>
<td>Medical Record Availability</td>
<td>Amount Dispensed</td>
<td>Facility</td>
<td>Diagnosis Code &amp; Type</td>
</tr>
<tr>
<td></td>
<td>Etc.</td>
<td>Etc.</td>
<td>Principal Discharge Diagnosis</td>
</tr>
<tr>
<td><strong>Registry Data</strong></td>
<td><strong>State Vaccine</strong></td>
<td><strong>Inpatient Pharmacy</strong></td>
<td><strong>Inpatient Transfusion</strong></td>
</tr>
<tr>
<td>Death</td>
<td>Patient ID</td>
<td>Patient ID</td>
<td>Patient ID</td>
</tr>
<tr>
<td>Cause of Death</td>
<td>Vaccination Date</td>
<td>Administration Date &amp; Time</td>
<td>Administration Start &amp; End Date &amp; Time</td>
</tr>
<tr>
<td>Source</td>
<td>Admission Date</td>
<td>Encounter ID</td>
<td>Encounter ID</td>
</tr>
<tr>
<td>Confidence</td>
<td>Vaccine Code &amp; Type</td>
<td>National Drug Code (NDC)</td>
<td>Transfusion Administration ID</td>
</tr>
<tr>
<td>Etc.</td>
<td>Provider</td>
<td>Route</td>
<td>Transfusion Product Code</td>
</tr>
<tr>
<td></td>
<td>Etc.</td>
<td>Dose</td>
<td>Blood Type</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Etc.</td>
<td>Etc.</td>
</tr>
</tbody>
</table>

#### Available Data Elements

https://www.sentinelinitiative.org/sentinel/data/distributed-database-common-data-model
Data Quality Review and Characterization Process

1. Preparation
   Sentinel Operations Center prepares quality review and characterization package for new ETL

2. Transformation
   Data Partner transforms source data into the Sentinel Common Data Model

3. Distribution
   Sentinel Operations Center distributes quality assurance package to Data Partners

4. Model Compliance
   Data Partner runs quality review and characterization package completing the following:
   - Level 1 checks
   - Level 2 checks
   Quality review and characterization package outputs list of errors or anomalies (flags) identified during data checks
   Data Partner resolves these flags and sends a detailed report to the Sentinel Operations Center

5. Review & Characterization
   Sentinel Operations Center receives output from Data Partner and reviews
   - Level 2 checks
   - Level 3 checks
   - Level 4 checks
   Sentinel Operations Center runs additional quality assurance checks:
   - Level 2 checks
   - Level 3 checks
   - Level 4 checks
   Sentinel Operations Center evaluates any additional flags and creates an issue report for Data Partner to address

6. Completion
   Data Partner investigates issues identified in report generated by the Sentinel Operations Center and resolves remaining flags

7. Approval
   Sentinel Operations Center Quality Assurance Manager approves ETL for use in queries

* On average, there are 44 flags identified by the program and 10 additional flags identified by the Sentinel Operations Center per ETL

https://www.sentinelinitiative.org/sentinel/data-quality-review-and-characterization
Data Quality Checks and Examples

**Level 1 Checks**
- **Completeness**
  - Admission date is not missing value
- **Validity**
  - Admission date is in date format

**Level 2 Checks**
- **Accuracy**
  - Admission date occurs before the patient’s discharge date
- **Integrity**
  - Admission date occurs within the patient’s active enrollment period

**Level 3 Checks**
- **Consistency of Trends**
  - There is no sizable percent change in admission date record counts by month-year

**Level 4 Checks**
- **Plausibility**
  - There is no sizable percent change in the number of prostate cancer encounters by sex*

*Under development

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
Session I: Establishing a High-Quality RWD Ecosystem

Join the conversation with #RWE2019
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.
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
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?

<table>
<thead>
<tr>
<th>Selected lab-related data checks (failure criteria)</th>
<th>% of DataMarts passing (most recent refresh; n=41)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than 80% of lab results mapped to LOINC</td>
<td>85%</td>
</tr>
<tr>
<td>Less than 80% of quantitative lab results specify the normal range</td>
<td>42%</td>
</tr>
<tr>
<td>Less than 80% of quantitative lab results mapped to LOINC specify specimen source &amp; result unit</td>
<td>37%</td>
</tr>
</tbody>
</table>
Curation as a learning process – data latency

Latency / completeness of data

- Developed latency calculation & incorporated into data curation

Questions:
- “How complete & up-to-date are the data?” (DSMB)
- “What’s the data censoring date for participants?” (Statistician)

No events? Or just no data?
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 a 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.
Session II: Curating and Assessing Fit-for-Use RWD Derived from Electronic Health Records
Lunch

Join the conversation with #RWE2019
Session III: Leveraging Digital Technology for Patient-Generated Health Data
Understanding PGHD Data Quality in the Real World

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

ерамirez@evidation.com
@ерамirez

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

OCTOBER 3, 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
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.
PGHD allows for measuring novel outcomes for chronic conditions at the population level.

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

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

*p < 0.05  **p < 0.001; FDR ADJUSTED

**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

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.

113 participants

- 31 symptomatic participants
- 82 Healthy Control (HC)
- 24 Mild Cognitive Impairment (MCI)
- 7 Mild Alzheimer’s Disease Dementia (AD)
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 behaviogram 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. FOURZANJANI 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?
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.
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 is designed to support

Discovery
Clinical Development
Post Approval
Diverse research cohorts for cutting edge research

Adult Quality of Care Program

Adult Translational Research Study

Pediatric Risk Stratification Study

Online Patient Survey Study
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

**SECONDARY RWD**

IBD SmartForm

*IBD PLEXUS*
Information Management Lifecycle

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

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

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

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

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

Prep-to-research tools
- Data querying capabilities
- Data visualization and insight tools
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
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
Powering IBD Plexus

- Patient & Clinician Engagement Platform
- Biobank & Lab Information Management System
- Central Reference Labs
- Data & Analytic Platforms
- High Performance Computing
- Researcher Portal
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 of 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
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
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?
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”

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


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

Edwards JK¹, Hoo PT¹, Stürmer T¹.

Author information

1 Department of Epidemiology, 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

Non-Active Comparator
New User Design: Confounding by Indication (Obesity)

Obese → Insulin

Obese → Glargine → CRC
Obese → NPH → CRC

Normal weight → No insulin

T2 Diabetes

Active-Comparator New User Design: No Confounding by Indication (Obesity)
### Confounding Control by Design: BMI

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

<table>
<thead>
<tr>
<th>BMI (kg/m²), n (%)</th>
<th>Glargine</th>
<th>NPH</th>
</tr>
</thead>
<tbody>
<tr>
<td>(&lt;19)</td>
<td>4 (0.7)</td>
<td>8 (1.9)</td>
</tr>
<tr>
<td>19 to (&lt;25)</td>
<td>77 (13.4)</td>
<td>67 (16.3)</td>
</tr>
<tr>
<td>25 to (&lt;30)</td>
<td>150 (26.1)</td>
<td>105 (25.5)</td>
</tr>
<tr>
<td>30 to (&lt;35)</td>
<td>146 (25.4)</td>
<td>104 (25.2)</td>
</tr>
<tr>
<td>35 to (&lt;40)</td>
<td>114 (19.9)</td>
<td>64 (15.5)</td>
</tr>
<tr>
<td>40 to (&lt;45)</td>
<td>45 (7.8)</td>
<td>36 (8.7)</td>
</tr>
<tr>
<td>(\geq 45)</td>
<td>38 (6.6)</td>
<td>28 (6.8)</td>
</tr>
</tbody>
</table>

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.”

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
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

Two-year RD: -0.73% (95% CI: -1.40%, -0.06%)

Initial Treatment

Two-year RD: 0.44% (95% CI: -0.22%, 1.09%)

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
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 $X \rightarrow C_M$ arrow, adjusting for $C_M$ cannot generate bias
  - Will partially control for $C$
- If there is an $X \rightarrow C_M$ arrow but no $C \rightarrow X$ 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%

C more common in RWD

C more common in trial

- <50%
- 50-80%
- >80%

Michael Webster-Clark, PharmD, PhD, unpublished
Bias in Stratum C=1 When Specificity is 0.99

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

C more common in RWD

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

P(C)=2.5%  P(C)=50%  P(C)=97.5%

C more common in trial

>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
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
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
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.”

---

Causal frameworks are needed to actually replicate the RCTs

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
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

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

Observational study conducted concurrently with RCT:
A high profile case where RCTs and observational studies differed

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


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


And subsequent studies revealed the reasons why.


Was randomization the issue?

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

Expert panel

- Duke-Margolis Center for Health Policy
- Eli Lilly & Company
- GlaxoSmithKline
- Food and Drug Administration
- ISPOR
- National Pharmaceutical Council
- ...and more

Sponsors

Research partners

- Amgen
- AstraZeneca
- Merck
- Optum
- Pfizer
- Sanofi
- UCB BioSciences, Inc.
**OPERAND study design**

**Focus:** On-label effectiveness in defined subgroups

<table>
<thead>
<tr>
<th><strong>Number of teams and trials</strong></th>
<th>Two academic institutions will independently replicate two identical target trials:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1. ROCKET for atrial fibrillation</td>
</tr>
<tr>
<td></td>
<td>2. Lead-2 for Type 2 diabetes control</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Data</strong></th>
<th>(a) Claims data alone and (b) Claims + EHR, each used for sensitivity analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Data will be restricted to inclusion and exclusion criteria of pivotal RCT</td>
</tr>
<tr>
<td></td>
<td>and on-label indication</td>
</tr>
</tbody>
</table>

| **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   |

<table>
<thead>
<tr>
<th><strong>Approach</strong></th>
<th>To ensure comparability, the teams will:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Be given a common clinical question and the study RCT protocol</td>
</tr>
<tr>
<td></td>
<td>• Be given defined set of anticipated methods</td>
</tr>
<tr>
<td></td>
<td>• Have flexibility to use their own methods in certain areas</td>
</tr>
<tr>
<td></td>
<td>• Initially, be restricted to inclusion/exclusion criteria</td>
</tr>
</tbody>
</table>
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

<table>
<thead>
<tr>
<th>Many methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Classification trees</td>
</tr>
<tr>
<td>• Random forests</td>
</tr>
<tr>
<td>• Bagging and boosting models</td>
</tr>
<tr>
<td>• Ridge, lasso, and elastic net regression</td>
</tr>
<tr>
<td>• Support vector machines</td>
</tr>
<tr>
<td>• Ensembles</td>
</tr>
<tr>
<td>• Neural networks</td>
</tr>
<tr>
<td>• And many others…</td>
</tr>
</tbody>
</table>

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
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

And how we get to causal associations with RWE

- Randomized Controlled Trials
  - Random treatment assignment
  - Controlled outcome measurement
  - Easy to understand and communicate

Franklin, Glynn, Martin, Schneeweiss, CP&T 2019
Causal study designs: Contemplate the target trial

Parallel group RCT

Washout period w/o study drug use → Exposed

Comparator (or placebo)

Cohort study

Washout period w/o study drug use → Exposed

Comparator

= 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
Estrogen plus Progestin and the Risk of Coronary Heart Disease

JoAnn E. Manson, M.D., Dr.P.H., Judith Haas, M.D., Karen C. Johnson, M.D., M.P.H., Jacques E. Rossouw, M.D., Ann louise R. Assaf, Ph.D., Norma n 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., Eran 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)

Risk of CHD events
1.42 (0.92–2.20)

Risk of CHD events
0.60 (0.43–0.83)

Risk of stroke
0.66 (0.53–0.82)

Risk of stroke
0.75 (0.58–0.98)

Risk of stroke
5.79 (1.81 to 18.6)
<table>
<thead>
<tr>
<th></th>
<th>RCT</th>
<th>RW</th>
<th>RW</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Risk of death (any)</strong></td>
<td>0.78 (0.68–0.89)</td>
<td>0.79 (0.60–1.03)</td>
<td>0.62 (0.58–0.66)</td>
</tr>
<tr>
<td><strong>Risk of hip fracture</strong></td>
<td>1.05 (0.88–1.25)</td>
<td>1.02 (0.83–1.24)</td>
<td>0.48 (0.27–0.83)</td>
</tr>
</tbody>
</table>
4. Proceed if
   a) Outcome observable with specificity
   b) Sufficient outcome surveillance
   c) Sufficient patient similarity is reached

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
**RCT**

- Risk of death (any)
  - 0.87 (0.74–1.01)

**RW**

- Risk of death (any)
  - 0.80 (0.69–0.92)

**Immortal study time**

- 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

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
Line programming for healthcare database analytics

Lacks transparency
Lacks reproducibility against intended protocol

Nothing wrong with sharing programming code but is not helpful…

… as it does not clarify whether the indented study was implemented accurately
Making it easier for decision makers to fully understand RWE

Longitudinal design visualization

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. Wang1,2 | Sebastian Schneeweiss1,2 | Marc L. Berger3 | Jeffrey Brown4 | Frank de Vries5 | Ian Douglas6 | Joshua J. Gagne1,2 | Rosa Gini7 | Olaf Klungel8 | C. Daniel Mullins9 | Michael D. Nguyen10 | Jeremy A. Rassen11 | Liam Smeeth9 | Miriam Sturkenboom12
on behalf of the joint ISPE-ISPOR Special Task Force on Real World Evidence in Health Care Decision Making

<table>
<thead>
<tr>
<th>TABLE 2 Reporting specific parameters to increase reproducibility of database studies*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
</tr>
<tr>
<td>A. Reporting on data source should include:</td>
</tr>
<tr>
<td>A.1 Data provider</td>
</tr>
<tr>
<td>A.2 Data extraction date (DED)</td>
</tr>
<tr>
<td>A.3 Data sampling</td>
</tr>
<tr>
<td>A.4 Source data range (SDR)</td>
</tr>
<tr>
<td>D. Reporting on exposure definition should include:</td>
</tr>
<tr>
<td>D.1 Type of exposure</td>
</tr>
<tr>
<td>D.2 Exposure risk window (ERW)</td>
</tr>
<tr>
<td>D.2a Induction period†</td>
</tr>
<tr>
<td>D.2b Stockpiling‡</td>
</tr>
<tr>
<td>D.2c Bridging exposure episodes*</td>
</tr>
</tbody>
</table>

---

*ISPE International Society for Pharmacoeconomics and Outcomes Research

†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...
Replicating 150 database studies

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

* binary/categorical

Study ID

- Covariate codes not reported
- Covariate codes reported

No difference
Replicating 150 database studies

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

- 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

Correlation between effect sizes (publication vs replication)

- Correlation coefficient = 0.74
Effect Size and Confidence Limits

Correlation between effect sizes
(publication vs replication)

- Correlation coefficient = 0.74
A pathway with regulatory validation

---

**Sponsor implements analysis**

**Is setting adequate for RWD analysis?**

- Yes → **Is data quality fit for purpose?**
  - Yes → Statistical analysis plan → **Feasibility analysis***
  - No → RCT

- No → RCT

---

**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**.

---

**Regulator checks and re-analyses**

- Plan for additional analyses → Regulatory and HTA consideration

---

**Analysis** → **Structured reporting**

---

**Validated RWE analytics platform with audit trails**

---

**ClinicalTrials.gov registration**:

- TECOS – NCT03936062
- LEADER – NCT03936049
- CARMELINA – NCT03936036
- CANVAS – NCT03936010
- SAVOR-TIMI – NCT03936023

---

* Franklin, Glynn, Martin, Schneeweiss. CPT 2019
How well can RWD analyses reproduce RCT findings?

Phase 1
- Candidate RCTs
- Select target RCTs
- Set up scalable RWD analytics platform
- Reproduce RCTs with RWD

Phase 2
- RWD study infrastructure

Phase 3
- List of RCTs to be reproduced with RWD
- Document exclusions: Limited RWD, Key measurements missing, Extremely strong confounding etc.
- Scalable RWD infrastructure
- Expert group guidance

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

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

Risk of CV events (3P-MACE)

HR = 0.91 (0.79 – 1.05)

NCT03648424

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

followed by

RCT

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

Risk of CV events (3P-MACE)

HR = ???

NCT01243424

CAROLINA

2019 Harvard / Brigham Division of Pharmacoepidemiology

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

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

PI: Franklin, Schneeweiss
ADA June 10, 2019

CAROLINA

Patrano E. et al. Diab Care 2019;42: June 14 epub
Independent Evidence Dossiers for decision makers?

- **Relevance**: Does the study, as implemented, address the intended question? 
  - Yes

- **Validity**: Are the methods valid? (bias minimizing)
  - Yes

- **Replicability**: Are the results directly replicable? (in the same data source)
  - Yes

- **Robustness**: Are the results robust? (to investigator specifications)
  - Yes

- **Transportability**: Are the results transportable? (to other populations/data)
  - Yes
Session IV: Methodological and Analytical Considerations for Observational Studies
### Summary of Essential Characteristics of Adequate and Well-Controlled Investigations

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>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.</td>
</tr>
<tr>
<td>2</td>
<td>The study uses a design that permits valid comparison w/ a control to provide a quantitative assessment of drug effect.</td>
</tr>
<tr>
<td>3</td>
<td>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.</td>
</tr>
<tr>
<td>4</td>
<td>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.</td>
</tr>
<tr>
<td>5</td>
<td>Adequate measures are taken to minimize bias on the part of the subjects, observers, and analysts of the data.</td>
</tr>
<tr>
<td>6</td>
<td>The methods of assessment of subjects’ response are well-defined and reliable.</td>
</tr>
<tr>
<td>7</td>
<td>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.</td>
</tr>
</tbody>
</table>
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

---

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 Ascertained 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?

- Findings characterized the demographic and clinical characteristics of aNSCLC patients treated with immune checkpoint inhibitors and assessed the ability to generate real-world endpoints (rwOS, rwPFS, rwTTNT, rwTTD).

Key Conclusions:

- High level of shared characteristics among the varying data sets despite varying sample sizes, data capture processes, and data sources demonstrating the feasibility of identifying aNSCLC patients 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.

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

<table>
<thead>
<tr>
<th>Project Goals:</th>
<th>Explore potential endpoints that may be fit for assessing long term benefits of a product compared to an existing alternative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Focus</td>
<td>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?</td>
</tr>
</tbody>
</table>
| 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

- PD-(L)1 + Doublet Chemotherapy
- PD-(L)1 Monotherapy
- Platinum Doublet Chemotherapy

Percentage of Male aNSCLC Patients

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

- 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
Demographic and Clinical Characteristics

Median and Lower/Upper Quartiles of Age at Index

- PD-(L)1 + Doublet Chemotherapy
- PD-(L)1 Monotherapy
- Platinum Doublet Chemotherapy

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
Histology of Patients with aNSCLC Per Treatment Category

- Nonsquamous: 81.2%
- Squamous: 18.8%
Kaplan Meier Curves per Treatment Group

**rwOS**

Platinum Doublet Chemotherapy

PD-(L)1 Monotherapy

**rwTTD**

Platinum Doublet Chemotherapy

PD-(L)1 Monotherapy

PD-(L)1 + Doublet Chemotherapy
Estimates of Median Time per Treatment Category

rwOS

- PD-(L)1 + Doublet Chemotherapy
- PD-(L)1 Monotherapy
- Platinum Doublet Chemotherapy

rwTTD

- PD-(L)1 + Doublet Chemotherapy
- PD-(L)1 Monotherapy
- Platinum Doublet Chemotherapy
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 heterogeneous 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 Ascertained Endpoints in Routine Clinical Care Settings

Join the conversation with #RWE2019
Open Comment Period
Closing Remarks
Adjournment