Approaches to patient follow-up for clinical trials: What’s the right choice for your study?

Keith Marsolo, PhD

Instructor
Department of Population Health Sciences
Duke Clinical Research Institute
Duke University School of Medicine
Scenario

- Planning a pragmatic clinical trial that leverages real-world data for some / all of the data collection

- Some of the sites are part of a distributed research network, but it necessary to include others

- What approaches do you take for the remaining sites? How do you make sure you are not paying for more data than you need?
Caveats

- Definition of “real-world data” here limited to events / outcomes found in electronic health records (EHRs) and / or claims

- Focusing on Medicare claims; private payers out of scope (for now)

- Privacy-preserving record linkage is out of scope – any linkage that might be needed can happen at the study coordinating center
Factors to consider

- What question(s) are you trying to answer with the data?

- How do you align questions to available data sources?

- What are the capabilities of potential sites?
  - Support for different data delivery methods (report, database, etc.)
  - “Sophistication” of implementation

- What is the per-site budget allocation?
Not all questions are created equal (in terms of data required)

- Hospitalizations
  - Was the participant hospitalized in the past year?
  - Was the most recent hospitalization the result of heart failure?

- Laboratory results
  - What was the participant’s most recent eGFR value?
  - What were the participant’s Hematocrit values 2 years prior to enrollment?

- Medication usage
  - How long was the participant on Xolair?
  - What medications were they taking on March 1, 2015?
  - Do their treatment patterns reflect standard of care?
Data sources & data delivery (primarily EHR)

- Can the source be used to answer the question?

- For a given source, there may be multiple ways of delivering the data
  - Some delivery methods may have pre-defined views or summarizations of the data
  - Do these views provide the right level of detail?

- Some delivery methods implicitly assume a certain level of data standardization
  - If you intend to take advantage of that standardization, have you made sure that the sites are compliant?
Data standardization in the EHR

- Most health systems still do not natively generate/capture data in standard terminologies (e.g., SNOMED, LOINC, RxNORM, etc.)
- If delivery method utilizes a standard, need to understand what progress sites have made, if any, before use
- Example – mapping lab tests to LOINC
  - All tests or just a subset?
  - All results or just from a specific point in time?
- Depending on mapping, how you ask the question will influence results
  - All Hemoglobin A1c results
  - All results for LOINC codes 4548-4, 41995-2, and 17855-8
Sources of Data & Modes of Delivery

Source

EHR
Sources of Data & Modes of Delivery

- EHR
- Claims (CMS)
Sources of Data & Modes of Delivery

Source

EHR

Claims (CMS)

Participant
Sources of Data & Modes of Delivery

Who procures the data?

- EHR
- Claims (CMS)
- Participant
Sources of Data & Modes of Delivery

- **Who** procures the data?
  - Patient

- **Source**
  - EHR
  - Claims (CMS)
  - Participant
<table>
<thead>
<tr>
<th>Source</th>
<th>EHR</th>
<th>Claims (CMS)</th>
<th>Participant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinician /</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Staff</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Sources of Data & Modes of Delivery

Who procures the data?

- Patient
- Clinician / Staff
- IT / Data Expert

<table>
<thead>
<tr>
<th>Source</th>
<th>EHR</th>
<th>Claims (CMS)</th>
<th>Participant</th>
</tr>
</thead>
</table>

- EHR
- Claims (CMS)
- Participant

- Patient
- Clinician / Staff
- IT / Data Expert
Sources of Data & Modes of Delivery

Who procures the data?

<table>
<thead>
<tr>
<th>Source</th>
<th>EHR</th>
<th>Claims (CMS)</th>
<th>Participant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient</td>
<td>Blue Button / Direct (Summary of Care)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Apple Health Records (FHIR)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinician / Staff</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IT / Data Expert</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Sources of Data & Modes of Delivery

<table>
<thead>
<tr>
<th>Source</th>
<th>EHR</th>
<th>Claims (CMS)</th>
<th>Participant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient</td>
<td>Blue Button / Direct (Summary of Care)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinician / Staff</td>
<td>Apple Health Records (FHIR)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IT / Data Expert</td>
<td>Clinician-Generated Report</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Sources of Data & Modes of Delivery

Who procures the data?

- **Patient**
  - Blue Button / Direct (Summary of Care)
  - Apple Health Records (FHIR)

- **Clinician / Staff**
  - Clinician-Generated Report

- **IT / Data Expert**
  - Analyst-Generated Report
  - Database Extract
  - Common Data Model
  - Application Programming Interface (e.g., FHIR)

**EHR**

**Claims (CMS)**

**Participant**
### Sources of Data & Modes of Delivery

<table>
<thead>
<tr>
<th>Source</th>
<th>EHR</th>
<th>Claims (CMS)</th>
<th>Participant</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Who procures the data?</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient</td>
<td>Blue Button / Direct (Summary of Care)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinician / Staff</td>
<td>Apple Health Records (FHIR)</td>
<td>Blue Button 2.0</td>
<td></td>
</tr>
<tr>
<td>IT / Data Expert</td>
<td>Clinician-Generated Report</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Analyst-Generated Report</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Database Extract</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Common Data Model</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Application Programming Interface (e.g., FHIR)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Sources of Data & Modes of Delivery

**EHR**
- Blue Button / Direct (Summary of Care)
- Apple Health Records (FHIR)
- Clinician-Generated Report
- Analyst-Generated Report
- Database Extract
- Common Data Model
- Application Programming Interface (e.g., FHIR)

**Claims (CMS)**
- Blue Button 2.0

**Participant**
- Database Extract (Research Identifiable Files)

<table>
<thead>
<tr>
<th>Source</th>
<th>Patient</th>
<th>Clinician / Staff</th>
<th>IT / Data Expert</th>
</tr>
</thead>
<tbody>
<tr>
<td>Who procures the data?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Sources of Data & Modes of Delivery

**Who procures the data?**

- **Patient**
  - EHR
    - Blue Button / Direct (Summary of Care)
    - Apple Health Records (FHIR)
  - Claims (CMS)
    - Blue Button 2.0
  - Participant
    - Portal / Mobile App

- **Clinician / Staff**
  - EHR
    - Clinician-Generated Report
  - Claims (CMS)
    - Database Extract (Research Identifiable Files)

- **IT / Data Expert**
  - EHR
    - Analyst-Generated Report
    - Database Extract
    - Common Data Model
    - Application Programming Interface (e.g., FHIR)
## Sources of Data & Modes of Delivery

<table>
<thead>
<tr>
<th>Source</th>
<th>EHR</th>
<th>Claims (CMS)</th>
<th>Participant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient</td>
<td>Blue Button / Direct (Summary of Care)</td>
<td>Blue Button 2.0</td>
<td>Portal / Mobile App</td>
</tr>
<tr>
<td>Patient</td>
<td>Apple Health Records (FHIR)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinician / Staff</td>
<td>Clinician-Generated Report</td>
<td></td>
<td>Call Center</td>
</tr>
<tr>
<td>Clinician / Staff</td>
<td>Analyst-Generated Report</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IT / Data Expert</td>
<td>Database Extract (Research Identifiable Files)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IT / Data Expert</td>
<td>Common Data Model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IT / Data Expert</td>
<td>Application Programming Interface (e.g., FHIR)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Who procures the data?
- Patient
- Clinician / Staff
- IT / Data Expert
Modes of Delivery – EHRs
**Blue Button / Direct**

- Patient can request a structured (XML) Summary of Care document with information about most recent visit & some longitudinal values.

- **Pros:**
  - All patients can obtain from their EHR

- **Cons:**
  - Completeness of implementation varies by site/EHR
  - Text-based document
  - Typically needs to be brokered through an app (e.g., Hugo)
  - If care is received from multiple systems, need to request multiple documents

### Table 1: What to Include in Summary of Care Documents

<table>
<thead>
<tr>
<th>Include in Summary of Care Documents</th>
<th>Include in Summary of Care Documents</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Patient name</td>
<td>• Care plan field, including goals and instructions</td>
</tr>
<tr>
<td>• Referring or transitioning provider’s name and office contact information (EP only)</td>
<td>• Care team, including the primary care provider of record and any additional known care team members beyond the referring or transitioning provider and the receiving provider.</td>
</tr>
<tr>
<td>• Procedures</td>
<td>• Discharge instructions (Hospital Only)</td>
</tr>
<tr>
<td>• Encounter diagnosis</td>
<td>• Reason for referral (EP only)</td>
</tr>
<tr>
<td>• Immunizations</td>
<td></td>
</tr>
<tr>
<td>• Laboratory test results</td>
<td></td>
</tr>
<tr>
<td>• Vital signs (height, weight, blood pressure, BMI)</td>
<td></td>
</tr>
<tr>
<td>• Smoking status</td>
<td></td>
</tr>
<tr>
<td>• Functional status, including activities of daily living, cognitive and disability status</td>
<td></td>
</tr>
<tr>
<td>• Demographic information (preferred language, sex, race, ethnicity, date of birth)</td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Summary of care documents must also include:</em></td>
</tr>
<tr>
<td></td>
<td>• Current problem list (providers may also include historical problems at their discretion)</td>
</tr>
<tr>
<td></td>
<td>• Current medication list (a list of medications that a patient is currently taking)</td>
</tr>
<tr>
<td></td>
<td>• Current medication allergy list (a list of medications to which a given patient has known allergies)</td>
</tr>
</tbody>
</table>

Apple Health Records

- iPhone users have the ability to download records from their EHR(s) into their Health app
- More computable than Summary of Care document – discrete data, not just XML

Pros:
- Health app already installed (need secondary app for data sharing)
- Process to share results with other apps is easy
- Supported by ~210 health systems (and growing)

Cons:
- Leverages Fast Healthcare Interoperability Resources (FHIR) as a backend (not a bad thing)
  - However – need to understand the quality of the FHIR implementation – what’s available vs. the rest of the EHR
- Permissions allow patients to share all records, or “ask when updates available” – may result in loss over time
- Participants need to make a new connection for every health system in which they receive care
- App is in beta & no Android equivalent (for now)
Clinician-generated reports

- Most EHRs provide functionality that allows clinicians to generate on-demand reports geared towards answering care management-type questions (e.g., who received flu shot in last 30 days, who was in the ED last night, etc.)

- Pros:
  - Low-cost; can be generated in seconds
  - Real-time results

- Cons:
  - Limited ability to pull longitudinal results; geared towards “most recent” values – most recent lab result, date of last test
  - Clinicians may not know that they have the ability to do this – training & support varies by health system
Analyst-generated report / Database extract

- Work through local / vendor-based IT resources to generate a query from the site’s reporting database and/or data warehouse

- Pros:
  - “Lowest common denominator” approach for obtaining large-scale extracts
  - If pulling all/subsets of a database table or a standard format (e.g., Summary of Care), can often reuse the same query across vendors
  - Once implemented, sites can typically automate production & delivery

- Cons:
  - Approach may not be feasible for smaller sites or sites without local IT support
  - Complex queries rely on skillset/knowledge of local analyst – quality will vary across sites
  - Timeline / cost is variable
Common Data Model (CDM)

- Sites that participate in distributed research networks may have their data loaded into a CDM (e.g., PCORnet, Sentinel, OMOP/OHDSI, VDW, etc.)

- Pros:
  - Process to develop/distribute query is relatively straightforward
  - Can submit one query and get back results from the whole network
  - Most networks perform some level of data curation, though curation varies

- Cons:
  - Data elements of interest may not be in the CDM (not place in CDM and/or site has not loaded them)
  - Large studies will likely need to go beyond a single network
Application programming interface (e.g., FHIR)

- Standardized interface that allows data to be requested via discrete function calls

- Pros:
  - Allows for easier integration into apps or other programs
  - Can be used to pre-populate case report forms
  - If all sites have the same set of APIs, the “query” can be reused
  - Office of the National Coordinating for Health Information Technology (ONC) recently proposed a rule requiring all EHRs to support FHIR APIs as mechanism of data exchange

- Cons:
  - Potential for data mapping issues
  - A traditional query may translate into dozens/hundreds of API calls
  - Most sites have limited experience in delivering data this way
  - Skillset required to maintain/administer APIs is highly-specialized (i.e., sites have few of these people & they are very in-demand)
Modes of Delivery – CMS Claims
CMS Blue Button 2.0

- Contains four years of Medicare Part A, B and D data for 53 million Medicare beneficiaries in a discrete format (requested via FHIR API)

- Pros:
  - Can obtain data directly from the participant – avoid CMS file charges
  - Data should update as they are made available by CMS
  - CMS has recently proposed that all Medicare Advantage organizations, Medicaid managed care plans, CHIP managed care entities and Qualified Health Plan issuers in the Federally-Funded Exchanges support similar APIs

- Cons:
  - Need to go through a secondary app (i.e., cannot download to Apple Health)
  - Not all participants may follow through / continue to allow access for life of the study

CMS NPRM:


Image source:
https://bluebutton.cms.gov/
CMS Research Identifiable Files

- Traditional process of requesting CMS data through ResDAC
- Pros:
  - Data are well-curated
  - Complete data for all patients in Finder File
- Cons:
  - Need to go through CMS request process
  - Can be expensive
  - Latency may be a factor for some studies
Modes of Delivery – Participant-as-Source
Portal / Mobile app

- Participants self-report events / outcomes via a web portal or mobile app

- Pros:
  - Relatively low cost
  - Single solution for all sites
  - Increased participant engagement

- Cons:
  - Potential for recall bias for some events
  - May lose participants for longer studies
  - Some participants may not be comfortable/capable of using portal/app
Call center

- Staff can reach out to participants via phone (or text / messaging) to follow-up if portal/app assessments are not completed

- Pros:
  - Can mitigate some loss-to-follow-up
  - Some participants may prefer to interact with call center

- Cons:
  - Staffing costs are non-trivial
  - Certain demographics may be less interested in answering the phone (i.e., millennials)
Summary

- Trials with many sites will require a “patchwork quilt” of approaches (for now)
  - Quilt will look different depending on the needs of the trial

- Clinician-generated reports are an often-overlooked option

- Direct-from-patient solutions (i.e., Blue Button / Health Records) offer a relatively low-cost way of obtaining data on trial participants
  - Unlikely to obtain *all* data from *all* patients in this way
  - For some trials, that may not be a problem
  - If it is, need to consider site-based solution – otherwise, just wasting effort

- Regulations are moving industry towards more standardized methods of data exchange via APIs
  - Solves the data model problem (hooray!)
  - Until data are collected/generated using same standards/formats as the API, still need to understand the EHR-to-interface mapping
  - In particular, what is NOT available via the API