Data Sharing for a Learning Health System

The pSCANNER Clinical Data Research Network

Daniella Meeker
Mark Elson
Barbara Filkins
Agenda

- Clinical Data Research Networks
- pSCANNER
  - Overview
  - Data Architecture
  - Governance
- pSCANNER and HIE in CA
Clinical Data Research Networks

- Facilitate Comparative Effectiveness Research (CER) across large networked organizations
- Utilization of clinical data in EHRs and other electronic systems, and administrative data

PCORI / PCORnet

- Patient-Centered Outcomes Research Institute
- PCORnet: National network of 13 funded CDRNs and 20 Patient Powered Research Networks (PPRNs)
Products of a Learning Health System

- **Analysis models for causal inference and program evaluation**
  - Program X is 15% more effective at preventing readmission than Program Y
  - Drug A has a greater risk for adverse events than Drug B

- **Quality measurement reports**
  - Practice N is achieving 80% of quality indicators
  - Practice L is achieving 60% of quality indicators

- **Predictive models for patient-centered medicine**
  - For patients like John, Drug C is safer and more effective than Drug D.
  - For patients with Debbie’s goals and comorbidity profile, Program Y is better than Program X
Both facilitate data sharing

Similar range of technical architectures

- pSCANNER is “hybrid-federated” with edge servers at participating sites
- CDRNs can also be “centralized” with a Clinical Data Repository

Some similar governance issues given multi-stakeholder structure
Convergence supporting a Learning Health System (LHS)

- HIEs
  - HIE for treatment of Individuals
  - HIE data aggregation for analytics, population health, QI; reporting
  - Rapid-cycle, data-driven improvement

- CDRNs
  - Research for generation of new knowledge
  - Research feedback loops for analytics, population health, QI; reporting

LHS
Many HIEs more mature from an organizational, user adoption, and governance perspective

Greater use of standards for HIE than in data-sharing for research

Some CDRNs pursuing cutting-edge privacy-preserving technical approaches for analytics
30 Million People

UCSD is Lead Contractor

Primarily in California
Clinical Research Focus

- pSCANNER Research Focus
  - Congestive Heart Failure (CHF)
  - Obesity
  - Kawasaki disease

- Current Studies
  - Aspirin dosage for patients with heart disease
  - Bariatric surgery outcomes
  - Diabetes risk
  - Statin effectiveness study

- New Studies
  - Researchers may propose new studies / queries
Achieve interoperability between research and clinical care by 2019

- Patient Engagement
- Policymaker Engagement
  - Standards Committee participation
  - Contribute to data policy for interoperability, billing, and incentive programs
- Health System Engagement
  - Create data resources with business value
  - Create reporting resources with business value
- Research Engagement
  - Participate in multisite research standards development
  - Offer easy to use privacy-preserving infrastructure for multisite collaboration
pSCANNER Data Architecture
Health Economic Domains are not represented in other research and quality information models, but value was So Cal Health Leadership Priorities.
We just need to migrate the data from these systems to fit into that hole over there.

I’ll get the hammer.
Aside on Data Quality & Availability

Dx
Rx
Lab
Procedures

EHR
Finance
LIS

STANDARDIZED WAREHOUSE
Standardizing 11 Health Systems

CUSTOM PROGRAMMING

OMOP

QUALITY MODEL

OMOP

OMOP

OMOP

OMOP

OMOP

OMOP

OMOP

OMOP

OMOP

OMOP

SQL Server

QUALITY MODEL

ORACLE

QUALITY MODEL

QUALITY MODEL

QUALITY MODEL

QUALITY MODEL

QUALITY MODEL

QUALITY MODEL

QUALITY MODEL

QUALITY MODEL

QUALITY MODEL

QRDA

DISTRIBUTED ANALYTICS

NEXTGEN HEALTHCARE INFORMATION SYSTEMS
Aside Continued…

Dx
Rx
Lab
Procedures

EHR
Finance
LIS

CCD
QRDA

STANDARDIZED WAREHOUSE

We just need to migrate the data from these systems to fit into that hole over there. I’ll get the hammer.
Why do we need standards for computation if the data is already standardized?
Direct and Quantifiable Comparisons

dataset A

algorithm 1

algorithm 2

dataset A

30 day readmission

dataset A

dataset B

algorithm 1
Selecting a Standard for Computation Specification

- Sufficiently expressive to represent data processing concepts for **transactional, time-series data** (e.g. interval logic)
- Sufficiently expressive to represent data processing concepts that are **specific to healthcare** (e.g. time of administration, age of onset)
- Sufficiently expressive to represent **basic statistics** and **data analysis algorithms**
- **Supported/Adopted**

Diagram:

- **Quality Data Model Data Processing Semantics**
- **Predictive Model Markup Language**
- **SQL**
Disseminating LHS Products

- OMOP Reference Data Model
- Data Set Extraction Program
  - Extracted Data Set ("Flat File")
  - Analysis Program
    - Estimated Predictive Model
      - Predictive Model Computation
        - Patient Centered Prediction

- CCDC Record Processing Program
- Standardized Extracted Record

- Publish
- research
- care
- Patient Record

aka "Data Processing" "Computable Phenotype" "Cohort" "Inclusion Criteria" "Measure Denominator"
<table>
<thead>
<tr>
<th>Process we need to represent</th>
<th>Standard</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data processing rules</td>
<td>HQMF&gt;CQL</td>
<td>CMS, ONC, HL7 endorsed Part of EHR certification process New standards in trial use</td>
</tr>
<tr>
<td>Cohort definition rules</td>
<td>HQMF&gt;CQL</td>
<td>100’s of established data sets 1000s of cohort criteria</td>
</tr>
<tr>
<td>Data set description</td>
<td>QRDA PMML</td>
<td>QRDA – Quality Measurement EHR Certification &amp; CMS PMML – Data analysis</td>
</tr>
<tr>
<td>Data Analysis Methods</td>
<td>PMML</td>
<td>UCSD Data Mining Group Extensible to support model specifications</td>
</tr>
<tr>
<td>Data Analysis Results (Estimated Models, Produce Predictions)</td>
<td>PMML</td>
<td>Developed to represent results Adopted by most stats packages</td>
</tr>
<tr>
<td>Process Workflow</td>
<td>BPML?</td>
<td>TBD</td>
</tr>
</tbody>
</table>
This visualization was created through Academy Health’s EDM Forum, a project funded by the Agency for Healthcare Research and Quality, Grant #U18 HS02789. The tool is maintained by the University of Southern California.
Governance Objectives

- Define processes, roles, relationships
- Support IRB streamlining and coordination
- Develop policies & operational procedures
  - Consistent and transparent framework, able to harmonize with site P&Ps & align with PCORnet P&Ps
  - Process to create, test, and approve pSCANNER P&Ps
  - Specific cases: onboarding to pSCANNER, research request and approval process, informed consent
- Monitor/evaluate governance performance
Detail: Data Services

- Design (Template)
- Create (Populate)
- Manage
- Reporting
- Design
- Syntactic Alignment
- Semantic Interoperability
- Tokenization
- Masking
- Encryption
- Database Schema
- Standardization
- Quality
- DataSet Access
- Logging and Audit
- Fields to be Audited
- Review Processes
- Delete
- Append
- Update
- Creation
- Terminology
- Master Data Management
- Create (Populate)
- Update
- Append
- Delete
- Review Processes
- Fields to be Audited
- Logging and Audit
- DataSet Access
- Quality
- Reporting
- Manage
- Create (Populate)
- Design (Template)
- Design
- Syntactic Alignment
- Semantic Interoperability
- Tokenization
- Masking
- Encryption
- Database Schema
- Standardization
Governance Framework

Operational Elements
- Budget and Cost Model
- Privacy Impact Assessment
- BAA (where required)
- Data Use Agreement with Key elements
- Schedule Timeline
- Resource Plan as related to Budget and Cost Model

Site Boundary Conditions: Operational
- Operational Objectives
- Resource/Planning
- Implementation
- Compliance
- Patient/Participant Engagement

pSCANNER Research Governance

Policy Mainstays
- Technical Infrastructure
- Identity & Access Management
- Research Protocols
- Data Services

Policy Compliance Elements
- System Security Plan (NIST 800-53 def.)
- Access Control Model based on Data Classification
- Study Design
- Data Management Policies
  - Data Classification Policy
  - Informed Consent
- Data Protection Plan
- Data Risk Assessment

Resource/Planning
# Study Initiation and Approval

## pSCANNER Research Governance Workflow (DRAFT)

<table>
<thead>
<tr>
<th>Funding Proposal</th>
<th>Pre-Award</th>
<th>Post-Award</th>
<th>Site Authorization</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sponsor</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Determine the need for a multi-site research study</td>
<td>• Prepare proposal according to sponsor format</td>
<td>• Establish project management</td>
<td>• Post-Award Status Updates/Reports</td>
</tr>
<tr>
<td>• Select Principal Investigator (PI) and co-PI(s)</td>
<td>• Include site letter(s) of support</td>
<td>• Finalize protocol design and specification, data management plan</td>
<td></td>
</tr>
<tr>
<td>• Define patient recruitment options and sample size</td>
<td>• Develop budget model using inputs from review of pSCANNER study proposal</td>
<td>• Update pSCANNER Site Assessments</td>
<td></td>
</tr>
<tr>
<td><strong>PI</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Determine capability to participate in study in proposed time frame</td>
<td>• Review operational and policy compliance assessments</td>
<td>• Update status to sponsor</td>
<td>• PI Notified of Site Authorization</td>
</tr>
<tr>
<td>• Review participation requirements (training, reporting, etc.)</td>
<td>• Provide input to study proposal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Identify possible study sites and institutional delegates (site reps)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Prepare pSCANNER documents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>pSCANNER Coordinator</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• pSCANNER Study Request</td>
<td>• pSCANNER Review/Inputs</td>
<td>• pSCANNER Coordinator Notified of Site Authorization</td>
<td></td>
</tr>
<tr>
<td>• pSCANNER Site Assessment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ethics/IRB (pSCANNER/Site)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Determine capacity and time required to prepare ethics application</td>
<td>• Generate Ethics-Related Documents</td>
<td>• pSCANNER Site Assessment (Updated)</td>
<td></td>
</tr>
<tr>
<td>• Access and seek agreements to prepare needed documents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Site Representative</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Determine if study is consistent with site’s mission and contracts</td>
<td>• Generate Letter of Support</td>
<td>• Site Authorization</td>
<td></td>
</tr>
<tr>
<td>• Assess proposed budget and controls</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Assess availability of staff and other resources</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Assess availability of suitable type and number of patients</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**Notes:**
- **PI:** Principal Investigator
- **Sponsor:** Agency funding the study
- **Site Representative:** Site点的联系人
- **pSCANNER Coordinator:** Coordinating the study across multiple sites
- **Ethics/IRB:** Institutional Review Board responsible for ethical reviews and approvals
- **Funding Proposal:** Detailed proposal outlining the study's objectives, methods, and budget
- **Pre-Award:** Stage before funding is received
- **Post-Award:** Stage after funding is received
- **Site Authorization:** Approval and resources for site participation
- **pSCANNER Site Assessment:** Review process for site readiness
- **Generate Ethics-Related Documents:** Preparing documents necessary for ethical review
- **Generate Letter of Support:** Documentation to support site participation
- **Update Status to Sponsor:** Keeping the sponsor informed of study progress
- **Finalize IRB Submission:** Ensuring all documents are completed and submitted for review
- **Review final IRB submission:** Final review and approval by the IRB
- **Provide approval and documents to PI and Sponsor:** Sent to the PI and sponsor for final approval
- **PI Notified of Site Authorization:** Communication of site's approval to the PI

---

**Legend:**
- Blue arrows represent the flow of information and decision points.
- Orange arrows indicate the timeline and key stages of the workflow.
<table>
<thead>
<tr>
<th>Phase: Deliverable</th>
<th>Pre-Award</th>
<th>Post-Award</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harmonization</td>
<td>Letter of support</td>
<td>Must be verified and committed to prior to letter of support</td>
</tr>
<tr>
<td>• Study consistent with site mission and contracts?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Suitable types and number of patients available?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resource/Planning</td>
<td>Preliminary sizing estimates must be provided for study proposal to sponsor</td>
<td>Scope of work and final numbers must be agreed to before site authorization accepted by PI</td>
</tr>
<tr>
<td>• Budget</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Schedule Timeline</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Resource Availability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Implementation Metrics (Data-Related)</td>
<td>Scope of work must be stated and consistent with resource/planning numbers provided</td>
<td>All required agreements signed and executed by all parties</td>
</tr>
<tr>
<td>• Schedule/Timeline</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Level of Effort</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Appropriate data model / data inputs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Compliance</td>
<td>Alignment on policies must be verified via pSCANNER taxonomy OR site must agree to undergo pSCANNER policy compliance assessment</td>
<td>All required agreements signed and executed by all parties</td>
</tr>
<tr>
<td>• Policy Alignment based on review of policies involving site research, privacy, data classification, and others as required</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Did site undergo pSCANNER Policy Compliance Assessment?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• What were the findings?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Agreements (Needed)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Participant Agreement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• BAA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Data Use Agreement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient/Participant Engagement Required</td>
<td>Questions must be completed to determine need for actual patient contract</td>
<td>No change from pre-award analysis unless study protocol has been modified</td>
</tr>
<tr>
<td>• Aggregate data being requested? (No direct patient engagement, data-sharing agreement)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• DID or LDS being requested? (May need authorization, data sharing agreement)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Individually-identifiable data being provided? (Need to understand patient/participant engagement techniques, use of informed consent)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Governance Parallels

• HIE – 5 Domains
  – Stakeholder Engagement
  – Structure & Organization
  – Business & Finance
  – Privacy & Legal
  – Technology & Security

• CDRN -- Activities
  – Patient Engagement
  – Strategic Decision-Making
  – Contract Management
  – Policy and Procedure Management
  – IRB Coordination
  – Data Governance
  – Approval of Research Studies and Queries
  – Project Management
pSCANNER and HIE in CA
HIE Readiness for Research?

- Overwhelming HIE focus on treatment today
  - Pop health analytics emerging as secondary focus
- Research largely an unexplored domain
- Negative association of research with pharma and selling data, but...
- ... Emergence of CDRNs provides potential for large-scale, trusted partners
Questions for Discussion

- What policy and technical work would be required to enable researchers in CDRNs to access HIE data in California?

- What would be needed to justify the business case for HIEs?
  - Payment for study participation?
  - Ability for HIE participants to initiate research studies / queries?
  - Access to “research-grade” analytics platform for HIE users?
  - Consolidated quality reporting?

- Should CTEN explore implementing a transaction pattern for research?
Contact Information

Daniella Meeker, PhD
Assistant Professor, USC Keck School of Medicine
Director, Clinical Research Informatics
Southern California Clinical Translational Sciences Institute
dmeeker@usc.edu

Mark Elson, PhD
Principal, Intrepid Ascent
mark@intrepidascent.com

Barbara Filkins
Principal, Syntax2Semantics
filkins@impulse.net

pSCANNER website: http://pscanner.ucsd.edu/
PCORnet website: http://pcornet.org/