pSCANNER – A Framework To Facilitate Distributed Queries and Analytics from Healthcare Data Repositories

Raj N Manickam[1], Daniella Meeker[2], Michael Matheny[3], Xiaoqian Jiang[4], Vira Danak[1], Michael Hogarth[1], Lucila Ohno-Machado[4]
[1]University of California Davis Health Informatics, CA, [2]University of Southern California, CA, [3]Vanderbilt University Medical Center, Nashville, Tennessee,
[4]University of California San Diego, CA.

Objectives

- **pSCANNER**: PCORI's effort to build data networks and engage patients and delivery systems in sponsored research and translation
- pSCANNER will allow users to access data using pSCANNER portal without patient-level data leaving the institute
- **Study A**:

– Engage patients and clinicians in the network governance

– Prioritize research questions using Delphi consensus process

- **Study B**:

– Conduct patient surveys

– Identify and recruit patient cohorts

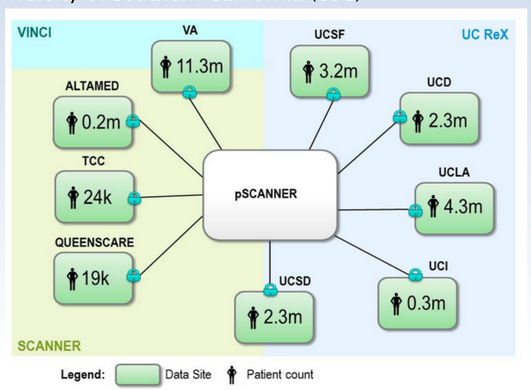
- Focus on 3 conditions:
  Obesity, Congestive Heart Failure, Kawasaki disease

Background

pSCANNER is designed to be a stakeholder-governed federated network that uses a distributed architecture to integrate data from three existing networks covering over 21 million patients in all 50 states:

pSCANNER participants

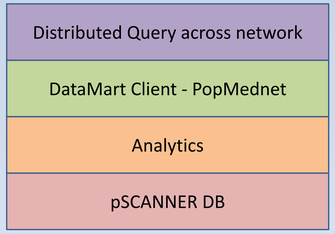
1. UC-ReX (UCD, UCLA, UCI, UCSF, UCSD)
2. VA (Salt Lake City, Tennessee Valley, San Diego)
3. S.Cal FQHCs – QueensCare, AltaMed, The Children's Clinic
4. University of Southern California (USC)



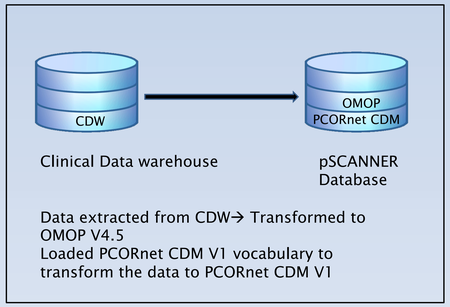
Methods and Results

Standard Data Model and Infrastructure

- All institutions in pSCANNER have agreed to standardize their data model to OMOP and to install a pSCANNER node to allow distributed computing, which greatly enhances distributed count query capabilities into multivariate analytics.



Components of pSCANNER node



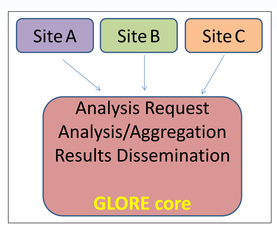
pSCANNER DB

- Use standard terminologies ICDx-CM, SNOMED-CT, HCPCS/CPT, RxNORM, LOINC
- Steps will be developed in standard format, using standard operating procedures and shared across participating institutions
- Additional tools for quality auditing and assessing validity and fitness for use in both research and other secondary uses of population-level data

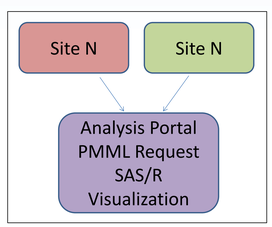
Privacy and Security Preservation

- pSCANNER will respond to PCORnet with results aggregated from all responding networks
- Each network has its own governance policies and processes for data sharing on the type of data requested (counts, statistics, others)
- Synchronous (real-time) and asynchronous (non real-time) modes of response will be supported

Analytical Framework



SCANNER Analytical node



pSCANNER Analytical node

Discussion

Potential Uses

- Distributed query network to benchmark cross institutional healthcare quality
- Health care providers in their local communities are empowered to respond proactively to disease outbreaks, understand the efficacy of drug treatments, and monitor health trends
- Access to large pool of data will enable to prioritize prevention procedures
- Healthcare research, example: researchers and providers could compare the effectiveness of different treatments and medications across different population without identifying individual information and data remaining behind the healthcare organization's firewalls

Future plans

- PCORI's Phase II goals for pSCANNER focuses on
  - Engaging patients, researchers, clinicians and health system in research and network governance
  - Preserving strong privacy and data security
  - Infrastructure that supports clinical trials within the network
  - Framework that fosters public trust in research
  - Sustainable research networks

References & Acknowledgements

pSCANNER: patient-centered Scalable National Network for Effectiveness Research. Ohno-Machado L, Agha Z, Bell DS, Dahm L, Day ME, Doctor JN, Gabriel D, Kahlon MK, Kim KK, Hogarth M, Matheny ME, Meeker D, Nebeker JR; *J Am Med Inform Assoc. 2014 Jul-Aug;21(4):621-6. doi: 10.1136/amiajnl-2014-002751. Epub 2014 Apr 29.*

Wu Y, Jiang X, Kim J, Ohno-Machado L. Grid Binary LOgistic REgression (GLORE): building shared models without sharing data. *Journal of the American Medical Informatics Association: JAMIA* 2012;19(5):758-764.

This research was supported by the Patient-Centered Outcomes Research Institute (PCORI) Contract CDRN-1306-04819