PopMedNet: Data Characterization Tools in Distributed Networks

Melanie Davies, HPHCI; Jessica Malenfant, MPH, HPHCI; Bruce Swan, LPP; Jeffrey Brown, PhD, HPHCI

HPHCI: Harvard Pilgrim Health Care Institute; LPP: Lincoln Peak Partners

Methods

• The PopMedNet and MSOC teams collaborated to identify high priority data characterization queries and define the requirements for the new querying functionality.
• A new data characterization model was defined and populated to enable querying.
• The teams assessed the requirements to ensure that the enhancements would meet the specific needs of MSOC while being general and extensible enough to be seamlessly implemented for other distributed research networks using PMN.
• Nine new PopMedNet Data Checker request types were designed, developed, and implemented within the Mini-Sentinel network. Users can run data checker queries using a point-and-click interface incorporated into the PMN platform. The nine request types focus on the following: Race, Ethnicity, Data completeness by table, Diagnosis codes, Discharge diagnosis types by encounter type, Presence of National Drug Codes, Pharmacy dispensing amount, Pharmacy dispensing days supply, Procedure codes, Query results can be displayed across all sites, within a specific site, or stratified by each site. Data are presented overall and by data partner in tables, graphs, and charts.

Conclusions

• The Data Checker tool allows networks to easily query commonly generated data characterization data (e.g., rate of missingness, proportions). This helps investigators better assess data availability before developing analytic code or distributing queries to data partners.
• The Data Checker tool and data model are generalizable for use in other distributed research networks. Any network that utilizes this new data characterization data model based on their own data checking and quality assurance processes can take advantage of this data checking tool in PopMedNet.

References