

## **Policy Tiger Team meetings:**

### **Discussion**

The primary focus for policy discussion relates to data collection standards and usage, including privacy and security... The APD use cases envision a broad range of uses which will require policies that address both identified and de-identified data. HIPAA currently includes definitions for de-identified data and also definition for a limited data set that can include some elements of identification that are necessary for certain kinds of research uses.

There is general support for utilizing the APCD proposed standard data set, but there may be policy issues relating to certain data elements (e.g., SSN).

Another important policy issue is collection of data from self-funded plans. The legislation authorizes DOH to include this information, but ERISA generally pre-empts states from imposing requirements on those plans. Vermont is currently testing the waters on this issue with one of their self-funded employers; Maine has adopted a voluntary data collection policy for these groups and about half of the self-funded plans participate. Insurers raised potential concerns related to APD use cases for insurance regulation.

The policy tiger team also reviewed different models and requirements relating to patient matching algorithms, including methodologies used by SPARCS and the proposed specifications for SHIN NY master patient index (MPI). We also discussed policy issues associated with enhanced demographic information including Affordable Care Act requirements to collect sex, race/ethnicity, primary language, and LGBT data collection.

### **Action Items**

- How do we define PHI and de-identified data?
- Review 18 data elements required by HIPAA to make data de-identifiable and outline direction to take on de-identifiable and identifiable data that will be needed in some cases.
- Follow up on issues in using SSN and collecting race, ethnicity, and LGBT data.
- Collection of LGBT data as per ACA
- Address issue of making data available for research
- Self-funded data collection issues
- Consult with ADK project and NYQA on how to use APD data for performance measurement.
- Follow up on NYS SSN Privacy Law

- Document nesting of levels of access and processes and agreements that will need to be in place for access to data.

### **Technical Tiger Team meetings:**

#### **Discussion**

The initial discussion was focused on identifying the standard datasets and existing data sources such as SPARCS, Fair Health etc. that are currently available. The National Cancer Institute's caBIG Grid provides another model for additional study regarding management of access and security of large amounts of cancer data for research purposes. There also was discussion around submission thresholds, payers expressing the need for an automated submission process and considerations relating to different technical architecture models. We identified additional technical experts from the participating stakeholders and had initial conversation about issues relating to strategies for dealing with very large data bases, what data models/architectures might be better suited than others, and operational issues such as data response time, frequency of data collection and historical storage of data.

#### **Action Items**

- Additional call with technical experts identified by members of the team.
- Technical Tiger Team members to provide comments on using the proposed NAHDO standard dataset.
- Requirements for data submission into the APD and for data that will be made available from the APD for analysis need to be identified clearly and distinctly. The Health Plan Association expressed that the requirements for data collection and release should be outlined separately. Plans would like clear direction on both data streams.
- Issues around defining a unique patient identifier for linking the different datasets is a key issue and would need to be addressed.